

Pred. No. is the number of results predicted by chance to have a

1
REFERENCE
AUTHORS
Baker, K. P., Ferrara, N., Gerber, H., Gerritsen, M. E., Goddard, A.,
Godowski, P. J., Gurney, A. L., Hillan, K. J., Marsters, S. A., Pan, J.,
Paoni, N. F., Stephan, J. P., Watanabe, C. K., Williams, P. M., Wood, W. I.

TITLE		and Ye, W.	
JOURNAL		Compositions and methods for the diagnosis and treatment of disorders involving angiogenesis	
		Patent: WO 0208284-A 215 31-JAN-2002; Genentech, Inc. (US) ; Baker, Kevin P. (US) ; Ferrara, Napoleone (US) ; Gerber, Hanspeter (US) ; Gerritsen, Mary E. (US) ; Goddard, Audrey (US) ; Godowski, Paul J. (US) ; Gurney, Austin L. (US) ; Hillan, Kenneth J. (US) ; Marsters, Scot A. (US) ; Pan, James (US) ; Paoni, Nicholas F. (US) ; Stephan, Jean-Philippe F. (US) ; Watanabe, Colin K. (US) ; Williams, P. Mickey (US) ; Wood, William I. (US)	
FEATURES		Location/Qualifiers	
source		1..2065	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
		Query Match 100.0%; Score 2065; DB 6; Length 2065; Best Local Similarity 100.0%; Pred. No. 0; Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GTGAATGTGAGGGTTTGATGACTTTTCAGATGCTTAGGAACACAGAGTGGGTGCAGGGGCC	60
Db	1	GTGAATGTGAGGGTTTGATGACTTTTCAGATGCTTAGGAACACAGAGTGGGTGCAGGGGCC	60
QY	61	CAGGACGGGCTGATTTCTTGGCGGAGGAGTAGGGTAAAGGTTCTGCATCAGCTCCTT	120
Db	61	CAGGACGGGCTGATTTCTTGGCGGAGGAGTAGGGTAAAGGTTCTGCATCAGCTCCTT	120
QY	121	AAAGGACAAAGGTAAACAGACCCAGACAGAGCTCGAGGGAGACTTTGACTTCAAGCCA	180
Db	121	AAAGGACAAAGGTAAACAGACCCAGACAGAGACTCGAGGGAGACTTTGACTTCAAGCCA	180
QY	181	CAGAAATGGTGAAGTGTGCGCGCGCCGCTCGCTCCTGCAGCGCTGTGCACCTA	240
Db	181	CAGAAATGGTGAAGTGTGCGCGCGCCGCTCGCTCCTGCAGCGCTGTGCACCTA	240
QY	241	GCGCTAGCACTTCCCGAGACACCGGATCCCGGGTAGGAGCGACCGCGGCGAGCACC	300
Db	241	GCGCTAGCACTTCCCGAGACACCGGATCCCGGGTAGGAGCGACCGCGGCGAGCACC	300
QY	301	AGCGCCAGCCGGCTCGGGTGCACACAGCGCTCACATGGGCTCGGGGCGCGGGCGGTG	360
Db	301	AGCGCCAGCCGGCTCGGGTGCACACAGCGCTCACATGGGCTCGGGGCGCGGGCGGTG	360
QY	361	TCCGCGGTCCGGCGCGTGTGCTGCTCCTCAGCTGCGCGGGCTGCCGCTCGGCGCAG	420
Db	361	TCCGCGGTCCGGCGCGTGTGCTGCTCCTCAGCTGCGGGGCTGCCGCTCGGCGCAG	420
QY	421	AACGACACGAGGCCATCGTGTGAGGGCAAGTGTCTGGTGTGCGACTCGAACCCG	480
Db	421	AACGACACGAGGCCATCGTGTGAGGGCAAGTGTCTGGTGTGCGACTCGAACCCG	480
QY	481	GCCACGAGCTCCAAAGGCTCTCTTCTCCCGCTGGGGATATCGGTCCGGGCGGCCAAC	540
Db	481	GCCACGAGCTCCAAAGGCTCTCTTCTCCCGCTGGGGATATCGGTCCGGGCGGCCAAC	540
QY	541	TCCAAAGTTCGGCGGTGCGGAGCACCAACACGAGCCATCCGAGATGAGCAAC	600
Db	541	TCCAAAGTTCGGCGGTGCGGAGCACCAACACGAGCCATCCGAGATGAGCAAC	600
QY	601	AAGACGCGCATCTTACTTCGATCAGATCCTGGTGAATGGGTAATTTTTTCACATTG	660
Db	601	AAGACGCGCATCTTACTTCGATCAGATCCTGGTGAATGGGTAATTTTTTCACATTG	660
QY	661	GAGTCTGCTTTGTAGCACCAAGAAAGGAATTTACAGTTTCAGTTTTCAGTGAATAAA	720
Db	661	GAGTCTGCTTTGTAGCACCAAGAAAGGAATTTACAGTTTTCAGTTTTCAGTGAATAAA	720
QY	721	GTCTACAGAGCCAACTATCCAGTTTAACTTGATGTTAAATGGAAAAACAGTAATATCT	780
Db	721	GTCTACAGAGCCAACTATCCAGTTTAACTTGATGTTAAATGGAAAAACAGTAATATCT	780

QY	781	GCCTTTGGGGGACAAAGATGTTTACTCGTGAAGCTGCCAAGATGGTGTCTGCTCTAC 840
Db	781	GCCTTTGGGGGACAAAGATGTTTACTCGTGAAGCTGCCAAGATGGTGTCTGCTCTAC 840
QY	841	CTAGATAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGTAAATTTGTTGGAGGCTGG 900
Db	841	CTAGATAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGTAAATTTGTTGGAGGCTGG 900
QY	901	CAGTATTCAGGTTTCTGGCTTCTCGTGTGTTCCCTTATAGGATCAATTTCTCATGA 960
Db	901	CAGTATTCAGGTTTCTGGCTTCTCGTGTGTTCCCTTATAGGATCAATTTCTCATGA 960
QY	961	TGTTTCATCCAGGTGAGGATGACCCACCTCCCTGAGTTATTTGGAAGATCAATTTTTCAT 1020
Db	961	TGTTTCATCCAGGTGAGGATGACCCACCTCCCTGAGTTATTTGGAAGATCAATTTTTCAT 1020
QY	1021	TGGATTTGATGCTTTTATTTGTTTCTCATGGGTGATATGATTTCTAAGGATTTCTAGCCT 1080
Db	1021	TGGATTTGATGCTTTTATTTGTTTCTCATGGGTGATATGATTTCTAAGGATTTCTAGCCT 1080
QY	1081	GTCTGAACCAATACAAAATTTTCACAGATTTTGTGTGTCTGTTCAGTATATTTTGA 1140
Db	1081	GTCTGAACCAATACAAAATTTTCACAGATTTTGTGTGTCTGTTCAGTATATTTTGA 1140
QY	1141	TTGGGACTCTAAGCAGATAATACCTATGCTTTAAATGTAAACAGTCAAAAGCTGTGCAAG 1200
Db	1141	TTGGGACTCTAAGCAGATAATACCTATGCTTTAAATGTAAACAGTCAAAAGCTGTGCAAG 1200
QY	1201	ACTTATTTCTGAATTTCTATTTCTCGGATTTCTAGTATTTACAGATGTGGAAATTTTAT 1260
Db	1201	ACTTATTTCTGAATTTCTATTTCTCGGATTTCTAGTATTTACAGATGTGGAAATTTTAT 1260
QY	1261	TGTTTAGTTTAAAGACTGGCAACAGCTTAAGGATTAAGAACTCTAAAGTTCTGAC 1320
Db	1261	TGTTTAGTTTAAAGACTGGCAACAGCTTAAGGATTAAGAACTCTAAAGTTCTGAC 1320
QY	1321	TTCAATCAACCGTTAGTGTGATATGCTGCAAGAACTGTATACTGTGTTTAAATATTTGAT 1380
Db	1321	TTCAATCAACCGTTAGTGTGATATGCTGCAAGAACTGTATACTGTGTTTAAATATTTGAT 1380
QY	1381	ATATTTGTTTATTTCTTTGGAATTTAGTTTGTGTTTCTGTTTAAAGAACTTGGATTT 1440
Db	1381	ATATTTGTTTATTTCTTTGGAATTTAGTTTGTGTTTCTGTTTAAAGAACTTGGATTT 1440
QY	1441	TTTTTTCAGTAACTGGTATTTATTTCTCTTAAATAAGTAAATGATGCTTGGCCAC 1500
Db	1441	TTTTTTCAGTAACTGGTATTTATTTCTCTTAAATAAGTAAATGATGCTTGGCCAC 1500
QY	1501	AAATTTTACCTTGACTAGATATCATCGACATGACTTCTCTCAAAAAAAGAAATGCTTCA 1560
Db	1501	AAATTTTACCTTGACTAGATATCATCGACATGACTTCTCTCAAAAAAAGAAATGCTTCA 1560
QY	1561	TAGTTGATTTTAAATTTGTAATGGAAGAGTCAATTTTCCAAAGTTATATTTTCTAAGA 1620
Db	1561	TAGTTGATTTTAAATTTGTAATGGAAGAGTCAATTTTCCAAAGTTATATTTTCTAAGA 1620
QY	1621	AGAAGATAGATCAATAATCTGCAAGCAAAAAAGTCTGTTACCCCAAACTTAAGTGCTCA 1680
Db	1621	AGAAGATAGATCAATAATCTGCAAGCAAAAAAGTCTGTTACCCCAAACTTAAGTGCTCA 1680
QY	1681	ATCCCTCAGGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATACTTTATTTGCTCAA 1740
Db	1681	ATCCCTCAGGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATACTTTATTTGCTCAA 1740
QY	1741	CTTTAAATTAATGATTAATAACCACTTTTAAATAAACCTTAAGGTTTTTTTTTTTTC 1800
Db	1741	CTTTAAATTAATGATTAATAACCACTTTTAAATAAACCTTAAGGTTTTTTTTTTTTC 1800
QY	1801	CGTAGACATGACCACTTTTAACTGTTGGTGGATGCTGTGTTTCTTAATATACCTAT 1860
Db	1801	CGTAGACATGACCACTTTTAACTGTTGGTGGATGCTGTGTTTCTTAATATACCTAT 1860


```
Db 1441 TTTTTCAGTAACGGTATATATGTTTCTCTTAAATAAGTAATAGTGGCTTGCCCA 1500
QY 1501 AAATTACCTTGACATACGATATCATCGACATGCTCTCTCAAAAAGAAAGTCTTCA 1560
Db 1501 AAATTACCTTGACATACGATATCATCGACATGCTCTCTCAAAAAGAAAGTCTTCA 1560
QY 1561 TAGTTGTATTTAAATGTAATGTGAAGAGTCAATATTTTCAAGTATATTTTCTAAG 1620
Db 1561 TAGTTGTATTTAAATGTAATGTGAAGAGTCAATATTTTCAAGTATATTTTCTAAG 1620
QY 1621 AGAAGAATAGATCAATAATCTGACAGGAAAGTCTTACCCAAATCTTAAGTGTCA 1680
Db 1621 AGAAGAATAGATCAATAATCTGACAGGAAAGTCTTACCCAAATCTTAAGTGTCA 1680
QY 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAATCTTATCTTTATGCTCAA 1740
Db 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAATCTTATCTTTATGCTCAA 1740
QY 1741 CTTTAAATTAAGTATGATAATACCACTTTATTAATAAACCCTAAGGTTTTTTTTTC 1800
Db 1741 CTTTAAATTAAGTATGATAATACCACTTTATTAATAAACCCTAAGGTTTTTTTTTC 1800
QY 1801 CGTAGACATGACCACTTTATTAACCTGGTGGGATGCTGTGTTCTAAATTAACCTAT 1860
Db 1801 CGTAGACATGACCACTTTATTAACCTGGTGGGATGCTGTGTTCTAAATTAACCTAT 1860
QY 1861 TTTTCAAGGCTCTCTGTTATTTGAAGTATCATCTGGTTTTGGCTTAATCTTTAAATTG 1920
Db 1861 TTTTCAAGGCTCTCTGTTATTTGAAGTATCATCTGGTTTTGGCTTAATCTTTAAATTG 1920
QY 1921 TATATATTATCTGTTAGCTAATTAATAATCAAAATCCAAATATCCAAATTAAGTCA 1980
Db 1921 TATATATTATCTGTTAGCTAATTAATAATCAAAATCCAAATATCCAAATTAAGTCA 1980
QY 1981 ATATCTGCTCTTTGTATAGTCAATGAAATTCATAAAATTAATTTATGCTGTATAGAA 2040
Db 1981 ATATCTGCTCTTTGTATAGTCAATGAAATTCATAAAATTAATTTATGCTGTATAGAA 2040
QY 2041 TAAAGATTAAATATGTTTAAAAAA 2065
Db 2041 TAAAGATTAAATATGTTTAAAAAA 2065

RESULT 3
AX697151
LOCUS AX697151 2065 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 219 from Patent WO0078961.
ACCESSION AX697151
VERSION AX697151.1 GI:29498104
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
Godowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,
Grimaldi,C.J., Hillan,K.J., Paoni,N.F., Roy,M.A. and Watanabe,C.K.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0078961-A 219 28-DEC-2000;
Genentech Inc.(US)
FEATURES
source location/Qualifiers
1..2065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2065; DB 6; Length 2065;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1081 GTCTGAACCAATACAAAATTTTCACAGATTATTTCTGTGTCCTCTTTCAGTATATTTGGA 1140
Db 1081 GTCTGAACCAATACAAAATTTTCACAGATTATTTCTGTGTCCTCTTTCAGTATATTTGGA 1140
QY 1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAATATGTAACAGTCAAAAGCTGCTGCAAG 1200
Db 1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAATATGTAACAGTCAAAAGCTGCTGCAAG 1200
QY 1201 ACTTATTTCTGAATTTCTCTTCTGGGATTTACTGTAATTTAGTTTACAGATGCTGGAATTTTATT 1260
Db 1201 ACTTATTTCTGAATTTCTCTTCTGGGATTTACTGTAATTTAGTTTACAGATGCTGGAATTTTATT 1260
QY 1261 TGTCTAGTTTAAAGACTGCGACACAGCTCTAAGGATAGAAATCTCTAAAGTCTGAC 1320
Db 1261 TGTCTAGTTTAAAGACTGCGACACAGCTCTAAGGATAGAAATCTCTAAAGTCTGAC 1320
QY 1321 TTCAATCAACGGTGTGATGATCTGCCAAGAACTGTATCTGTTGTTCTGTGTAATAATTTTATT 1380
Db 1321 TTCAATCAACGGTGTGATGATCTGCCAAGAACTGTATCTGTTGTTCTGTGTAATAATTTTATT 1380
QY 1381 ATATTTCTTTTATTTCTTTGGAATTTAGTTTGTGTTCTGTTGTAATAATTTTATT 1440
Db 1381 ATATTTCTTTTATTTCTTTGGAATTTAGTTTGTGTTCTGTTGTAATAATTTTATT 1440
QY 1441 TTTTTCAGTAACCTGTAATTTATGTTTCTCTTAAATAAGTAATGATGCTTGGCCAC 1500
Db 1441 TTTTTCAGTAACCTGTAATTTATGTTTCTCTTAAATAAGTAATGATGCTTGGCCAC 1500
QY 1501 AAATTTACCTTGACTACGATATCATGACATGACTCTCTCAAAAAGAAAGAAAGTCTCA 1560
Db 1501 AAATTTACCTTGACTACGATATCATGACATGACTCTCTCAAAAAGAAAGAAAGTCTCA 1560
QY 1561 TAGTTGTAATTTAATTTGATGATGTAAGAGTGCATATTTTCCAAAGTTATATTTTCTAAGA 1620
Db 1561 TAGTTGTAATTTAATTTGATGATGTAAGAGTGCATATTTTCCAAAGTTATATTTTCTAAGA 1620
QY 1621 AGAAGAAATAGATCATATAATCTGCAAGAAAGTCTGCTTACCCAAATCTAAGTGTCTCA 1680
Db 1621 AGAAGAAATAGATCATATAATCTGCAAGAAAGTCTGCTTACCCAAATCTAAGTGTCTCA 1680
QY 1681 ATCCCTGAGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATCTTATTTATGCTCAA 1740
Db 1681 ATCCCTGAGCTCAGCAAAAACAGCTCCCTCCGAGGGAATCTTATCTTATTTATGCTCAA 1740
QY 1741 CTTTAAATTAATGATGATAATAACACATTTATTTAAACCACTAAGGTTTTTTTTTTTTC 1800
Db 1741 CTTTAAATTAATGATGATAATAACACATTTATTTAAACCACTAAGGTTTTTTTTTTTTC 1800
QY 1801 CGTAGACATGACCACTTTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 1860
Db 1801 CGTAGACATGACCACTTTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 1860
QY 1861 TTTTCAAGCTTCTGTTGATTTGAAGTATCATCTGTTGTTTGTCTTAACTTAACTTAAATG 1920
Db 1861 TTTTCAAGCTTCTGTTGATTTGAAGTATCATCTGTTGTTTGTCTTAACTTAACTTAAATG 1920
QY 1921 TATATATTTATCTGTTTGTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 1980
Db 1921 TATATATTTATCTGTTTGTAGTATTAATTAATTAATTAATTAATTAATTAATTAATG 1980
QY 1981 ATATCTTGTCTTTGTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 1981 ATATCTTGTCTTTGTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
QY 2041 TAAAGATTATATATGTTTAAAAAAA 2065
Db 2041 TAAAGATTATATGTTTAAAAAAA 2065
```

RESULT 4
AY358527 2065 bp mRNA linear PRI 03-OCT-2003
LOCUS
DEFINITION Homo sapiens clone DNA6526 GSGR718 (UNQ718) mRNA, complete cds.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY358527
AY358527.1 GI:37182176
FLI_CDNA:
Homo sapiens (human)

REFERENCE
AUTHORS

1 (bases 1 to 2065)
Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J., Seehagiri S., Simmons L., Singh J., Smith V., Stinson J., Vasts A., Vandlen R., Watanabe C., Weand D., Woods K., Xie M.H., Yansura D., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I. and Godowski P.

TITLE

The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment
Genome Res. 13 (10), 2265-2270 (2003)

JOURNAL
PUBMED
REFERENCE

12975309
2. (bases 1 to 2065)
Clark H.F.

AUTHORS
JOURNAL

Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
Location/Qualifiers

FEATURES
source

1. 2065
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DNA6526"
1. 2065
/locus_tag="UNQ718"
337. 942
/locus_tag="UNQ718"
/notes="PRO1382"
/codon_start=1
/product="GSGR718"
/protein_id="AAQ88891.1"
/db_xref="GI:37182177"
/translation="MGSGRRALSAVPAVLVLTPLGLPVMAQNDTPIVLEKCLVVC
DGNPTDSSSSSSPLGISVRAASKVAFSAVRSTNHPSEMSKTRIIYFDQILNVV
GNFTLESVFVAPRKGIYSFPHVIVYQSQTIQVNLMLNGKPVISAFAGKRDVTEA
AINGVLLYLCKEDKIVLKLEKGLNGVGMQYSIFSGLVFLP"

gene

CDS

ORIGIN

Query Match 100.0%; Score 2065; DB 9; Length 2065;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGAATGTGAGGTTTGTGATCTTTCAGATGCTTAGGAACCCAGAGTGGGTGCAAGGGCCC 60
Db 1 GTGAATGTGAGGTTTGTGATCTTTCAGATGCTTAGGAACCCAGAGTGGGTGCAAGGGCCC 60
QY 61 CAGGCAAGGCTGATCTTTCGGCGGAGGAGTAGGTAAAGGTTCTTCATGAGTCCCTT 120
Db 61 CAGGCAAGGCTGATCTTTCGGCGGAGGAGTAGGTAAAGGTTCTTCATGAGTCCCTT 120
QY 121 AAAGGCAAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGAGCTTTCGACTTCAAGCCA 180
Db 121 AAAGGCAAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGAGCTTTCGACTTCAAGCCA 180
QY 181 CAGAAATGTGAAGTGTCTCGCGCGCCGCCCGCGCTCTCTGCAAGCTGTCGACCTA 240
Db 181 CAGAAATGTGAAGTGTCTCGCGCGCCGCCCGCGCTCTCTGCAAGCTGTCGACCTA 240
QY 241 GCGCTAGCATCTTCCGAGACCCGGGATCCCGGGGTAGGAGCGAGCGGGGAGCACC 300
Db 241 GCGCTAGCATCTTCCGAGACCCGGGATCCCGGGGTAGGAGCGAGCGGGGAGCACC 300
QY 301 AGCGCCAGCGGCTGCGGCTGCGCCCAACACGCTCACCATGGGCTCCGGGCGCGGGCGGCTG 360

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tx, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-885A10 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP5-885A10. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP5-885A10 is at 1 in this sequence. The true left end of clone RP11-380D15 is at 11755 in this sequence. The true right end of clone RP11-458E15 is at 27524 in this sequence.

FEATURES

source	Location/Qualifiers
repeat_region	1..117854 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="20" /clone="RP5-885A10" /clone_lib="RPCI-5" 585..740 /note="78 copies 2 mer aa 59% conserved"
repeat_region	594..649 /note="7 copies 8 mer aagaaag 76% conserved"
repeat_region	600..643 /note="11 copies 4 mer gga 79% conserved"
misc_feature	761..1245 /note="match: STS: Em:AA868507" complement(1551..7478) /gene="dJ885A10.1"
gene	complement(join(<1551..2061,4038..4154,7188..>7478))
mRNA	/gene="dJ885A10.1" /product="dJ885A10.1 (similar to cerebellin precursor)" /note="match: cDNAs: Em:M58583 match: ESTs: Em:AV72278" /evidence=not_experimental complement(join(1864..2061,4038..4154,7188..7478))
CDS	/gene="dJ885A10.1" /note="match: proteins: Sw:P98087 Sw:P23435" /codon_start=1 /evidence=not_experimental /product="dJ885A10.1 (similar to cerebellin precursor)" /protein_id="CAB90173.1" /db_xref="GI:7708438" /db_xref="SWISS-PROT:Q9NTU7" translation="MSGGRALSAVPAVLVLPLGLPWAQNDTPVLVKGKLVVC DNPATDKSSSPGIGIVRAANSKVAFAVSTNHEPSEMSKNKRIYDFDQILVNV GNFTLVSEVPARKYISFHFVIKYQQTQVNLMLNGKPVISAFAGDKDVTREA ATNGVLVLDKEDKVLKLEKGLVGVGQWSTPSGFLVFL" 2642..2958 /note="AluJo repeat: matches 1..306 of consensus" complement(2788..3347) /gene="dJ885A10.1" /note="match: GSS: Em:AQ530830" complement(2972..3470) /gene="dJ885A10.1" /note="match: GSS: Em:AQ317752" complement(3040..3495)
repeat_region	3341..3840 /note="match: GSS: Em:AQ408825" /note="match: GSS: Em:AQ514309" 3499..4061 /note="match: STS: Em:HS267E11T" 3608..3929 /note="match: GSS: Em:AQ594830" 4245..4280 /note="9 copies 4 mer ctcc 91% conserved" 4282..4325 /note="11 copies 4 mer tctt 81% conserved" complement(6245..6743) /gene="dJ885A10.1" /note="match: GSS: Em:B70819" 6729..7220 /note="match: GSS: Em:AQ676409" 6855..8929 /note="CpG island" /evidence=not_experimental
repeat_region	9008..9049 /note="7 copies 6 mer aaaaat 81% conserved"
repeat_region	9009..9052 /note="22 copies 2 mer aa 79% conserved"
repeat_region	9770..9892 /note="FLAM A repeat: matches 12..108 of consensus"
repeat_region	10273..10302 /note="15 copies 2 mer aa 96% conserved"
repeat_region	12095..12694 /note="LIP4 repeat: matches 5538..6140 of consensus"
repeat_region	13449..13623 /note="5 copies 35 mer 77% conserved"
repeat_region	13484..13599 /note="29 copies 4 mer tctc 75% conserved"
repeat_region	14187..14481 /note="AluY repeat: matches 1..294 of consensus"
repeat_region	15140..15329 /note="MIR repeat: matches 62..262 of consensus"
repeat_region	16419..16484 /note="MIR repeat: matches 124..192 of consensus"
misc_feature	16438..17076 /note="match: GSS: Em:AQ391974" 16830..16899 /note="2 copies 35 mer 92% conserved"
repeat_region	17052..17111 /note="L2 repeat: matches 2651..2710 of consensus"
repeat_region	17910..17987 /note="39 copies 2 mer tt 64% conserved"
repeat_region	19009..19305 /note="AluX repeat: matches 1..298 of consensus"
repeat_region	19337..19686 /note="THE1B repeat: matches 1..364 of consensus"
repeat_region	20347..20640 /note="AluX repeat: matches 1..294 of consensus"
repeat_region	20747..20874 /note="MER5B repeat: matches 1..145 of consensus"
repeat_region	22332..22267 /note="6 copies 6 mer cacaca 86% conserved"
repeat_region	22233..22264 /note="8 copies 4 mer acac 90% conserved"
misc_feature	complement(22555..23071) /note="match: GSS: Em:AQ451152" complement(22573..23081) /note="match: GSS: Em:AQ053765" 22757..23455 /note="L2 repeat: matches 1992..2710 of consensus"
repeat_region	24097..24357 /note="AluSg repeat: matches 1..261 of consensus"
misc_feature	complement(24374..24608) /note="match: GSS: Em:AQ677215" 24394..24844 /note="match: GSS: Em:AQ762377" 24839..25272


```

misc_feature      /note="match: GSS: Em:B58489"
24845..25393
misc_feature      /note="match: GSS: Em:AQ899089"
24942..25279
repeat_region     /note="match: GSS: Em:AQ222882"
25892..26161
repeat_region     /note="L2 repeat: matches 2090..2384 of consensus"
26231..26402
repeat_region     /note="MER3 repeat: matches 5..176 of consensus"
26407..27166
misc_feature      /note="LIPAL3 repeat: matches 5388..6155 of consensus"
complement(27157..27515)
repeat_region     /note="match: GSS: Em:AQ582450"
27175..27459
repeat_region     /note="AluX repeat: matches 1..285 of consensus"
27585..27738
repeat_region     /note="AluY repeat: matches 147..300 of consensus"
27744..27934
repeat_region     /note="L2 repeat: matches 2555..2750 of consensus"
28859..29029
repeat_region     /note="MIR repeat: matches 1..218 of consensus"
complement(29154..29702)
misc_feature      /note="match: GSS: Em:AQ385339"
29329..29553
repeat_region     /note="MIR repeat: matches 32..262 of consensus"
29593..30173
misc_feature      /note="match: GSS: Em:AQ052168"
match: SFS: Em:AQ052168"
29658..29737
repeat_region     /note="MER96 repeat: matches 1..71 of consensus"

```

```

Query Match      63.8%; Score 1318.4; DB 9; Length 117854;
Best Local Similarity 99.9%; Pred.No. 28-222;
Matches 1319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 740 TCAGGTTAACTGATGTTAAATGGAACCAAGTAATATCTCCCTTTCGGGGGACAAAG 799
DB 2066 TCTAGGTTAACTGATGTTAAATGGAACCAAGTAATATCTCCCTTTCGGGGGACAAAG 2007
QY 800 ATGTTACTCGTGAAGTCCAGCAATGGTCTCTCTACCTAGATAAGAGGATAAGG 859
DB 2066 ATGTTACTCGTGAAGTCCAGCAATGGTCTCTCTACCTAGATAAGAGGATAAGG 1947
QY 860 TTACTCTAAACTGAGAAAGTAAATTTGGTGGAGGCTGGCAGTATTCACAGGTGAGGCA 979
DB 1946 TTACTCTAAACTGAGAAAGTAAATTTGGTGGAGGCTGGCAGTATTCACAGGTGAGGCA 919
QY 920 GCTTTCGGTGTCCCTCATAGGATTCAAATTCCTCCATGATGTTTCATCCAGGTGAGGCA 1887
DB 1886 GCTTTCGGTGTCCCTCATAGGATTCAAATTCCTCCATGATGTTTCATCCAGGTGAGGCA 1827
QY 980 TGACCCACTCTCGAGTATTGGAGATCATTTTTTCATCATGATTGATGCTTTTATT 1039
DB 1826 TGACCCACTCTCGAGTATTGGAGATCATTTTTTCATCATGATTGATGCTTTTATT 1767
QY 1040 GGTTCCTCATGGTGGATATGATTCCTAAGGATTCAGGCTGCTGACCAATCAAAAT 1099
DB 1766 GGTTCCTCATGGTGGATATGATTCCTAAGGATTCAGGCTGCTGACCAATCAAAAT 1707
QY 1100 TTCACAGTATTATGCTGTGCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATA 1159
DB 1706 TTCACAGTATTATGCTGTGCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATA 1647
QY 1160 ATACCTAGCTTAATGTAACAGTCAAAAGCTGCTGCAAGCTTATTCGAAATTCATT 1219
DB 1646 ATACCTAGCTTAATGTAACAGTCAAAAGCTGCTGCAAGCTTATTCGAAATTCATT 1587
QY 1220 TCCTGGGATTACTGAATAGTATACAGATGCGAATTTTATTTGTTAGTTTAAAGACT 1279
DB 1586 TCCTGGGATTACTGAATAGTATACAGATGCGAATTTTATTTGTTAGTTTAAAGACT 1527
QY 1280 GGCACACCGGTCTAAGGATTAGAAACTCTAAAGTCTCTGACTTCAATCAACGGTTAGTGT 1339

```

```

DB 1526 GGCACACCGGTCTAAGGATTAGAAACCTTAAGTTCGACTTCAATCAACGGTTAGTGT 1467
QY 1340 GATACCTGCCAAGAACTGTATACCTGTGTAAATATATATGATTAATATTTGTTTATTCCTT 1399
DB 1466 GATACCTGCCAAGAACTGTATACCTGTGTAAATATATATGATTAATATTTGTTTATTCCTT 1407
QY 1400 TCGAATAGTATTGTTGTTCTTTGTAATAAAACCTTCGATTTTTCAGTAACTGGTAT 1459
DB 1406 TCGAATAGTATTGTTGTTCTTTGTAATAAAACCTTCGATTTTTCAGTAACTGGTAT 1347
QY 1460 TATGTTTCTCTTAAAAATAAGGTAAATGAATGCTTCCCAAAATTTACCTTGACTACGA 1519
DB 1346 TATGTTTCTCTTAAAAATAAGGTAAATGAATGCTTCCCAAAATTTACCTTGACTACGA 1287
QY 1520 TATCATCAGCATGACTCTCTCAAAAAAAGAAAGCTTCATAGTGTATTTTAAATGTA 1579
DB 1286 TATCATCAGCATGACTCTCTCAAAAAAAGAAAGCTTCATAGTGTATTTTAAATGTA 1227
QY 1580 TATGTGAAGAGTCAATATTTTCCAAAGTTATATTTTCTAAGAAAGAAATAGATCAATAAT 1639
DB 1226 TATGTGAAGAGTCAATATTTTCCAAAGTTATATTTTCTAAGAAAGAAATAGATCAATAAT 1167
QY 1640 CTGACACGAAAGAAAGTTCCTTACCCAAATCTAAGTGTCTCAATCCCTCAGCCTCAGCAA 1699
DB 1166 CTGACACGAAAGAAAGTTCCTTACCCAAATCTAAGTGTCTCAATCCCTCAGCCTCAGCAA 1107
QY 1700 ACAGCTCCCTCCGAGGAAATCTTATACCTTATTTGCTCACTTAAATTAATGATTTGA 1759
DB 1106 ACAGCTCCCTCCGAGGAAATCTTATACCTTATTTGCTCACTTAAATTAATGATTTGA 1047
QY 1760 TAAATAACCACTTTTATTAATAAACCTTAAGGTTTTTTTTTTCCTAGACATGACCACTTTA 1819
DB 1046 TAAATAACCACTTTTATTAATAAACCTTAAGGTTTTTTTTTTCCTAGACATGACCACTTTA 987
QY 1820 TTAAGTGTGGTGGAGTCTGTTGTTCTTAATATACCTATTTTCAAGGCTCTCTGTTGT 1879
DB 986 TTAAGTGTGGTGGAGTCTGTTGTTCTTAATATACCTATTTTCAAGGCTCTCTGTTGT 927
QY 1880 ATTTGAAGTATCATCTGTTTTCCTTAACCTTTTAAATGATATATATATCTGTTAG 1939
DB 926 ATTTGAAGTATCATCTGTTTTCCTTAACCTTTTAAATGATATATATCTGTTAG 867
QY 1940 CTAATATTAATTAATCAATATCCCATATCTAAATTTAGTGAATATCTGTTTGTGATA 1999
DB 866 CTAATATTAATTAATCAATATCCCATATCTAAATTTAGTGAATATCTGTTTGTGATA 807
QY 2000 GGTCTATGATTAATCAAAATTTATGCTGTGTATAGATAAAGATTAATATATGTTA 2059
DB 806 GGTCTATGATTAATCAAAATTTATGCTGTGTATAGATAAAGATTAATATATGTTA 747

RESULT 6
BC047076
LOCUS Homo sapiens, clone IMAGE:5314390, mRNA.
ACCESSION BC047076
VERSION BC047076.1 GI:28436739
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 926)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

```


cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsatne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Loulseged, H.,
Kowis, C. R., Sneed, A. G., Martin, R. G., Muzny, D. M., Nanavati,
A. N., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 91 Row: o Column: 23.

FEATURES

Location/Qualifiers
1..926
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones IMAGE:5314390
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 42.3%; Score 873; DB 9; Length 926;
Best Local Similarity 99.8%; Pred. No. 4.3e-144;
Matches 895; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 1169 CTTAAATGTAACAGTCAAAAGCTGCTGCAAGACTTATTCGAATTCATTCCTGGGAT 1228
DB 1 CTTAAATGTAACAGTCAAAAGCTGCTGCAAGACTTATTCGAATTCATTCCTGGGAT 60
QY 1229 TACTGAATAGTACAGATGGGAATTTATTTGTTTATTTTAAAGACTGGCAACGAG 1288
DB 61 TACTGAATAGTACAGATGGGAATTTATTTGTTTATTTTAAAGACTGGCAACGAG 120
QY 1289 GCTTAAGATTAGAAAACCTTAAAGCTTCTGACTTCAATCAAGGTTAGTGATCTGCC 1348
DB 121 GCTTAAGATTAGAAAACCTTAAAGCTTCTGACTTCAATCAAGGTTAGTGATCTGCC 180
QY 1349 AAAGAACTGTACTGTGTGTTAATATATGATTTATTTGTTTATTTCTTTGGAAATTAG 1408
DB 181 AAAGAACTGTACTGTGTGTT-ATAATGATTTATTTGTTTATTTCTTTGGAAATTAG 239
QY 1409 TTTGTTTGGTTCTTGTAATAAAACCTTGGATTTTTTTTTCAGTAACTGGTATTATGTTTC 1468
DB 240 TTTGTTTGGTTCTTGTAATAAAACCTTGGATTTTTTTTTCAGTAACTGGTATTATGTTTC 299
QY 1469 TCTTAAATTAAGTAAATGAATGCTTGGCCCAAAATTTACCTGACTAGTATCATCGA 1528
DB 300 TCTTAAATTAAGTAAATGAATGCTTGGCCCAAAATTTACCTGACTAGTATCATCGA 359
QY 1529 CATGACTTCTCTCAAAAAAAGAAAGTCTTCATAGTTGTTATTTAATTTGATATGTGAAA 1588
DB 360 CATGACTTCTCTCAAAAAAAGAAAGTCTTCATAGTTGTTATTTAATTTGATATGTGAAA 419
QY 1589 GAGTCATATTTTCCAAAGTTATATTTCTTAAGAGAGAAATAGATCATATAATCTGCAAGG 1648
DB 420 GAGTCATATTTTCCAAAGTTATATTTCTTAAGAGAGAAATAGATCATATAATCTGCAAGG 479
QY 1649 AAAAAAGTCTCTTACCCAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAACAGCTCCC 1708
DB 480 AAAAAAGTCTCTTACCCAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAACAGCTCCC 539
QY 1709 CTCGAGGGAATCTTATATCTTTATTTGCTCAACTTTAATTTAAATGATTAATAACCA 1768
DB 540 CTCGAGGGAATCTTATATCTTTATTTGCTCAACTTTAATTTAATGATTAATAACCA 599
QY 1769 CTTTATTAACCACTTAAGGTTTTTTTTTTTTCGCTGAGATGACCACTTTATTAATCTGT 1828

DB 600 CTTTATTAAAAAACCTAAGG-TTTTTTTTTTCCGTAGACATGACCACTTTATTAACTGGT 658
QY 1829 GGTGGGATGCTGTGTTTCTTAATATATACCTATTTTCAAGGCTTCTGTGTTATTGAAGT 1888
DB 659 GGTGGGATGCTGTGTTTCTTAATATATACCTATTTTCAAGGCTTCTGTGTTATTGAAGT 718
QY 1889 ATCATCTGGTTTTCCTTAACCTCTTTAAATTTGATATATTTATCTGTTTACCTAATATTA 1948
DB 719 ATCATCTGGTTTTCCTTAACCTCTTTAAATTTGATATATTTATCTGTTTACCTAATATTA 778
QY 1949 AATTCAAATATCCCATATCTAAATTTAGTGAATATCTGTGCTTTTGTATAGGTCATATG 2008
DB 779 AATTCAAATATCCCATATCTAAATTTAGTGAATATCTGTGCTTTTGTATAGGTCATATG 838
QY 2009 AATTCAAATATTTATTTATGCTCTATAGATAAAGATTAATATATATGTTAAAAAAA 2065
DB 839 AATTCAAATATTTATTTATGCTCTATAGATAAAGATTAATATATGTTAAAAAAA 895

RESULT 7

AC015860/c
LOCUS AC015860 248469 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-42A1, LOW-PASS SEQUENCE SAMPLING.
AC015860
AC015860.3 GI:9157673
VERSION HTG; HTGS PHASE0.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 248469)
Birken, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone RP11-42A1
Unpublished
2 (bases 1 to 248469)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., DeAngelis, M., Collins, S., Collamore, A.,
Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gage, D.,
Galland, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karakas, A., Klein, J.,
Lehoczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N.,
McSwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meidrim, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., VC, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi.6553969.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I613
Center clone name: 42_A_1

* NOTE: This record contains 221 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1	863:	contig of 863 bp in length
864	963:	gap of 100 bp
964	1841:	contig of 878 bp in length
1842	1941:	gap of 100 bp
1942	2831:	contig of 890 bp in length
2832	2931:	gap of 100 bp
2932	3755:	contig of 824 bp in length
3756	3855:	gap of 100 bp
3856	4750:	contig of 895 bp in length
4751	4850:	gap of 100 bp
4851	5744:	contig of 894 bp in length
5745	5844:	gap of 100 bp
5845	6713:	contig of 869 bp in length
6714	6813:	gap of 100 bp
6814	7715:	contig of 902 bp in length
7716	7815:	gap of 100 bp
7816	8692:	contig of 877 bp in length
8693	8792:	gap of 100 bp
8793	9674:	contig of 882 bp in length
9675	9774:	gap of 100 bp
9775	10656:	contig of 882 bp in length
10657	10756:	gap of 100 bp
10757	11630:	contig of 874 bp in length
11631	11730:	gap of 100 bp
11731	12610:	contig of 880 bp in length
12611	12710:	gap of 100 bp
12711	13622:	contig of 912 bp in length
13623	13722:	gap of 100 bp
13723	14592:	contig of 870 bp in length
14593	14692:	gap of 100 bp
14693	15567:	contig of 875 bp in length
15568	15667:	gap of 100 bp
15669	16420:	contig of 753 bp in length
16421	16520:	gap of 100 bp
16521	17444:	contig of 924 bp in length
17445	17544:	gap of 100 bp
17545	18422:	contig of 878 bp in length
18423	18522:	gap of 100 bp
18523	19463:	contig of 941 bp in length
19464	19563:	gap of 100 bp
19564	20472:	contig of 909 bp in length
20473	20572:	gap of 100 bp
20573	21444:	contig of 872 bp in length
21445	21544:	gap of 100 bp
21545	22434:	contig of 890 bp in length
22435	22534:	gap of 100 bp
22535	23414:	contig of 880 bp in length
23415	24356:	contig of 842 bp in length
24357	24456:	gap of 100 bp
24457	25329:	contig of 873 bp in length
25330	25429:	gap of 100 bp
25430	26315:	contig of 886 bp in length
26316	26415:	gap of 100 bp
26416	27322:	contig of 907 bp in length
27323	27422:	gap of 100 bp
27423	28291:	contig of 869 bp in length
28292	28391:	gap of 100 bp
28392	29255:	contig of 864 bp in length
29256	29355:	gap of 100 bp
29356	30234:	contig of 879 bp in length
30235	30334:	gap of 100 bp
30335	31224:	contig of 890 bp in length
31225	31324:	gap of 100 bp
31325	32255:	contig of 931 bp in length
32256	32355:	gap of 100 bp
32356	33282:	contig of 927 bp in length
33283	33382:	gap of 100 bp
33383	34259:	contig of 877 bp in length
34260	34359:	gap of 100 bp

34360	35252:	contig of 893 bp in length
35253	35352:	gap of 100 bp
35353	36214:	contig of 862 bp in length
36215	36314:	gap of 100 bp
36315	37222:	contig of 908 bp in length
37223	37322:	gap of 100 bp
37323	38320:	contig of 998 bp in length
38321	38420:	gap of 100 bp
38421	39306:	contig of 886 bp in length
39307	39406:	gap of 100 bp
39407	40276:	contig of 870 bp in length
40277	40376:	gap of 100 bp
40377	41271:	contig of 895 bp in length
41272	41371:	gap of 100 bp
41372	42255:	contig of 884 bp in length
42256	42355:	gap of 100 bp
42356	43225:	contig of 870 bp in length
43226	43325:	gap of 100 bp
43326	44186:	contig of 861 bp in length
44187	44286:	gap of 100 bp
44287	45182:	contig of 896 bp in length
45183	45282:	gap of 100 bp
45283	46191:	contig of 909 bp in length
46192	46291:	gap of 100 bp
46292	47171:	contig of 880 bp in length
47172	47271:	gap of 100 bp
47272	48137:	contig of 866 bp in length
48138	48237:	gap of 100 bp
48239	49176:	contig of 939 bp in length
49177	49276:	gap of 100 bp
49277	50214:	contig of 938 bp in length
50215	50314:	gap of 100 bp
50315	51073:	contig of 759 bp in length
51074	51173:	gap of 100 bp
51174	52022:	contig of 849 bp in length
52023	52122:	gap of 100 bp
52123	53012:	contig of 890 bp in length
53013	53112:	gap of 100 bp
53113	54013:	contig of 901 bp in length
54014	54113:	gap of 100 bp
54114	55000:	contig of 887 bp in length
55001	55100:	gap of 100 bp
55101	55972:	contig of 872 bp in length
55973	56072:	gap of 100 bp
56073	56948:	contig of 876 bp in length
56949	57048:	gap of 100 bp
57049	57945:	contig of 897 bp in length
57946	58045:	gap of 100 bp
58046	58929:	contig of 884 bp in length
58930	59029:	gap of 100 bp
59030	59927:	contig of 898 bp in length
59928	60027:	gap of 100 bp
60028	60922:	contig of 895 bp in length
60923	61022:	gap of 100 bp
61023	61918:	contig of 896 bp in length
61919	62018:	gap of 100 bp
62019	62930:	contig of 912 bp in length
62931	63030:	gap of 100 bp
63031	63978:	contig of 948 bp in length
63979	64078:	gap of 100 bp
64079	64999:	contig of 921 bp in length
65000	65099:	gap of 100 bp
65099	66010:	contig of 911 bp in length
66011	66110:	gap of 100 bp
66111	67066:	contig of 956 bp in length
67067	67166:	gap of 100 bp
67167	68147:	contig of 981 bp in length
68148	68247:	gap of 100 bp
68248	69185:	contig of 938 bp in length
69186	69285:	gap of 100 bp
69286	70218:	contig of 933 bp in length
70219	70318:	gap of 100 bp
70319	71216:	contig of 898 bp in length


```
* 13830 13929: gap of 100 bp
* 13930 15248: contig of 1319 bp in length
* 15249 15348: gap of 100 bp
* 15349 16757: contig of 1409 bp in length
* 16758 16857: gap of 100 bp
* 16858 18403: contig of 1546 bp in length
* 18404 18503: gap of 100 bp
* 18504 19522: contig of 1019 bp in length
* 19523 19622: gap of 100 bp
* 19623 21123: contig of 1501 bp in length
* 21124 21223: gap of 100 bp
* 21224 22242: contig of 1019 bp in length
* 22243 22342: gap of 100 bp
* 22343 23634: contig of 1292 bp in length
* 23635 23734: gap of 100 bp
* 23735 25389: contig of 1655 bp in length
* 25390 25489: gap of 100 bp
* 25490 26697: contig of 1208 bp in length
* 26698 26797: gap of 100 bp
* 26798 28096: contig of 1299 bp in length
* 28097 28196: gap of 100 bp
* 28197 29328: contig of 1132 bp in length
* 29329 29428: gap of 100 bp
* 29429 30679: contig of 1251 bp in length
* 30680 30779: gap of 100 bp
* 30780 32106: contig of 1327 bp in length
* 32107 32206: gap of 100 bp
* 32207 33847: contig of 1641 bp in length
* 33848 33947: gap of 100 bp
* 33948 35798: contig of 1851 bp in length
* 35799 35898: gap of 100 bp
* 35899 37908: contig of 2010 bp in length
* 37909 38008: gap of 100 bp
* 38009 39061: contig of 1053 bp in length
* 39062 39161: gap of 100 bp
* 39162 40468: contig of 1307 bp in length
* 40469 40568: gap of 100 bp
* 40569 42884: contig of 2716 bp in length
* 42885 43284: gap of 100 bp
* 43285 44719: contig of 1335 bp in length
* 44720 44819: gap of 100 bp
* 44820 46467: contig of 1648 bp in length
* 46468 46567: gap of 100 bp
* 46568 48871: contig of 2304 bp in length
* 48872 48971: gap of 100 bp
* 48972 50204: contig of 1233 bp in length
* 50205 50304: gap of 100 bp
* 50305 52244: contig of 1940 bp in length
* 52245 52344: gap of 100 bp
* 52345 53401: contig of 1057 bp in length
* 53402 53501: gap of 100 bp
* 53502 55091: contig of 1590 bp in length
* 55092 55191: gap of 100 bp
* 55192 58084: contig of 2893 bp in length
* 58085 58184: gap of 100 bp
* 58185 60261: contig of 2077 bp in length
* 60262 60361: gap of 100 bp
* 60362 61645: contig of 1284 bp in length
* 61646 61745: gap of 100 bp
* 61746 63437: contig of 1692 bp in length
* 63438 63537: gap of 100 bp
* 63538 64887: contig of 1250 bp in length
* 64888 64987: gap of 100 bp
* 64989 68773: contig of 3886 bp in length
* 68774 73467: gap of 100 bp
* 73468 73567: contig of 4594 bp in length
* 73568 77240: contig of 3673 bp in length
* 77241 77341: gap of 100 bp
* 77342 81090: contig of 3750 bp in length
* 81091 81190: gap of 100 bp
* 81191 85511: contig of 4321 bp in length
* 85512 85611: gap of 100 bp
```

```
* 85612 90487: contig of 4876 bp in length
* 90488 90587: gap of 100 bp
* 90588 95292: contig of 4705 bp in length
* 95293 95393: gap of 100 bp
* 95393 100087: contig of 4695 bp in length
* 100088 100187: gap of 100 bp
* 100188 106286: contig of 6099 bp in length
* 106287 106387: gap of 100 bp
* 106388 112851: contig of 6465 bp in length
* 112852 112951: gap of 100 bp
* 112952 117689: contig of 4738 bp in length
* 117690 117790: gap of 100 bp
* 117790 124437: contig of 6648 bp in length
* 124438 124537: gap of 100 bp
* 124538 133578: contig of 9041 bp in length
* 133579 133679: gap of 100 bp
* 133679 141606: contig of 7928 bp in length
* 141607 154145: contig of 12439 bp in length
* 154146 154245: gap of 100 bp
* 154246 165714: contig of 11469 bp in length
* 165715 165815: gap of 100 bp
* 165815 183553: contig of 17739 bp in length.
```

FEATURES

```
source
1. .183553
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="RP11-28A1"
Human Male BAC"
1. .1318
/note="assembly_fragment"
1419. .2757
/note="assembly_fragment"
2858. .4221
/note="assembly_fragment"
4322. .5593
/note="assembly_fragment"
5694. .6731
/note="assembly_fragment"
6832. .8028
/note="assembly_fragment"
8129. .9540
/note="assembly_fragment"
9641. .11147
/note="assembly_fragment"
11248. .12587
/note="assembly_fragment"
12688. .13829
/note="assembly_fragment"
13930. .15248
/note="assembly_fragment"
15349. .16757
/note="assembly_fragment"
16858. .18403
```

```
Query Match 29.8%; Score 615.6; DB 2; Length 183553;
Best Local Similarity 98.8%; Pred. No. 3e-98; Indels 0; Gaps 0;
Matches 621; Conservative 0; Mismatches 9;
```

```
Qy 1 GTGAATGTGAGGGTTTGATGACTTTCAGATGCTTAGGAACACAGAGTGGGTGAGGGGGCCC 60
Db 103843 GTCAACGACAGGGGGTTTCATGACTTTCAGATGCTTAGGAACACAGAGTGGGTGAGGGGGCCC 103902
Qy 61 CAGCGAGGGCTGATTTCTTGGCGGAGAGAGTAGGTAAGGGTTTCGATGAGTCCCTT 120
Db 103903 CAGCGAGGGCTGATTTCTTGGCGGAGAGAGTAGGTAAGGGTTTCGATGAGTCCCTT 103962
Qy 121 AAAGGCAAAAGGTAACAGAGCCAGAGAGCTCGAGGGGAGACTTTCAGTTCAAGCCA 180
Db 103963 AAAGGCAAAAGGTAACAGAGCCAGAGAGCTCGAGGGGAGACTTTCAGTTCAAGCCA 104022
Qy 181 CAGAAATTGGTGAAGTGTGCGCGCCGCCGCTGCTCTGCAGCGTGTGCGACCTA 240
```

```

RESULT 9
AX3117844
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

```

```

        614 bp      DNA      linear      PAT 14-DEC-2001
Sequence 19 from Patent WO0188135.
AX3117844
AX3117844.1 GI:17900720
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Padigaru,M., Grosse,W.M., Szerkeres,E.D., Casman,S.J.,
Ausbrook,J.P., Burgess,C.E., Taylor,S., Vernet,C.A.,
Baumgartner,J.C., Tschernev,V.T., Spytek,K.A., Li,L.,
Kekuda,R., Gangolli,E.A., Stone,D.J., Smithson,G. and
Maddougall,J.R.
G protein-coupled receptor proteins and nucleic acids encoding same
Patent: WO 0188135-A 19 22-NOV-2001;
Curagen Corporation (US)
Location/Qualifiers
1..614
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Query Match 29.7%; Score 614; DB 6; Length 614;
Best Local Similarity 100.0%; Pred. No. 2.2e-98;
Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 ACCATGGGCTCCGGCGCGCGGCGGTGTCGCGGGTCCGCGCGTCTGCTGCTCTCAG 393
|||||
1 ACCATGGGCTCCGGCGCGCGGCGGTGTCGCGGGTCCGCGCGTCTGCTGCTCTCAG 60
|||||
394 CTGCGGGGGCTCCCGTCTGGGCACAGAACACACGAGCCCCATCGTGTCTGAGGCGCAG 453
|||||
61 CTGCGGGGGCTCCCGTCTGGGCACAGAACACACGAGCCCCATCGTGTCTGAGGCGCAG 120
|||||
454 TGTCTGGTGGTGGCATCTCGAACCCCGGCCACGCACTCCAGGGTCTCTTTCCTCCCGC 513
|||||

```

[illegible]

LOCUS AC117551 256869 bp DNA linear HTG 03-MAR-2003
DEFINITION Mus musculus clone RP23-80N8, WORKING DRAFT SEQUENCE, 23 unordered
pieces.
ACCESSION AC117551.5 GI:28631357
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 256869)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 256869)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,I.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lebecky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,D., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 256869)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,I., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cooke,P., Cook,A., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 3, 2003 this sequence version replaced gi:28570472.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L17585
Center clone name: 80_N_8
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 248973 bases at least Q40
Consensus quality: 252607 bases at least Q30
Consensus quality: 253987 bases at least Q20
Insert size: 240000; agarose-fp
Insert size: 254669; sum-of-contigs
Quality coverage: 9.5 in Q20 bases; agarose-fp
Quality coverage: 9.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 5623: contig of 5629 bp in length
* 1
* 5630: gap of 100 bp
* 5723: contig of 725 bp in length
* 6454: gap of 100 bp
* 6455: contig of 1105 bp in length
* 6555: gap of 100 bp
* 7759: contig of 1581 bp in length
* 9341: gap of 100 bp
* 9441: contig of 1572 bp in length
* 11012: gap of 100 bp
* 11013: contig of 1815 bp in length
* 12927: gap of 100 bp
* 13027: contig of 1431 bp in length
* 14458: gap of 100 bp
* 14559: contig of 1925 bp in length
* 16483: gap of 100 bp
* 16583: contig of 1528 bp in length
* 18111: gap of 100 bp
* 18211: contig of 2125 bp in length
* 20336: gap of 100 bp
* 20337: contig of 2883 bp in length
* 23119: gap of 100 bp
* 23420: contig of 1505 bp in length
* 24925: gap of 100 bp
* 25025: contig of 3309 bp in length
* 28333: gap of 100 bp
* 28334: contig of 6314 bp in length
* 34747: gap of 100 bp
* 34748: contig of 7445 bp in length
* 34848: gap of 100 bp
* 42292: contig of 12525 bp in length
* 42993: gap of 100 bp
* 54917: contig of 100 bp
* 55017: gap of 100 bp
* 55018: contig of 13862 bp in length
* 74879: gap of 100 bp
* 74880: contig of 15472 bp in length
* 90451: gap of 100 bp
* 90452: contig of 17922 bp in length
* 108473: gap of 100 bp
* 108474: contig of 14457 bp in length
* 123030: gap of 100 bp
* 123031: contig of 54150 bp in length
* 177280: gap of 100 bp
* 177281: contig of 37963 bp in length
* 215343: gap of 100 bp
* 215344: contig of 41426 bp in length.
* 256869: contig of 41426 bp in length.
* 256869: Location/Qualifiers
* 1.256869
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"

```
/clone="RP23-80N8"
/clone_lib="RPCI-23 Female Mouse BAC"
1..5629
/note="assembly_fragment"
vector_end:SP6
vector_side:left
5730..6454
/note="assembly_fragment"
6555..7659
/note="assembly_fragment"
7760..9340
/note="assembly_fragment"
9441..11012
/note="assembly_fragment"
11113..12927
/note="assembly_fragment"
13028..14458
/note="assembly_fragment"
14559..16483
/note="assembly_fragment"
16584..18111
/note="assembly_fragment"
18212..20336
/note="assembly_fragment"
20437..23319
/note="assembly_fragment"
23420..24924
/note="assembly_fragment"
25025..28333
/note="assembly_fragment"
28434..34747
/note="assembly_fragment"
34848..42292
/note="assembly_fragment"
42393..54917
/note="assembly_fragment"
55018..74879
/note="assembly_fragment"
74980..90451
/note="assembly_fragment"
90552..108473
/note="assembly_fragment"
108574..123030
/note="assembly_fragment"
123131..177280
/note="assembly_fragment"
177381..215343
/note="assembly_fragment"
215444..256869
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN
Query Match 29.4%; Score 608; DB 2; Length 256869;
Best Local Similarity 73.6%; Pred. No. 7e-97;
Matches 1009; Conservative 0; Mismatches 300; Indels 61; Gaps 16;
QY 737 CTATCCAGTTAACTTGATGCTTAAATGGAAACAGTAATATCTGCTTTGGGGGACA 796
DB 247240 CTTTCCAGGTAACTTGATGCTTAAATGGAAACAGTAATCTCTGCAATTTGCTGGTGA 247181
QY 797 AAGATGTTACTCTGAGCTGCCACGAATGGTCTCTACCTACCTAGATAAAGAGGATA 856
DB 247180 AAGATGTTACCTGGAGCGCCACTAATGGAGTGTCTCTGACCTGGCAAGAGAGATA 247121
QY 857 AGGTTTACCTAAATCGGAAAGGTAATTTGGTTGGAGCTGGCAGTATTCACCGTTTT 916
DB 247120 AGGTTTACCTAAATCGGAAAGGTAATTTGGTTGGAGCTGGCAGTATTCACCGTTTT 247061
QY 917 CTGGCTTTCTGGTGTCTCCCTTAGGATTCATTTCTCATGATGTTCTATCCAGGTGAG 976
DB 247060 CTGGCTTTCTGGTGTCTCCCTTAGGATTCATTTCTCATGATGTTCTATCCAGGTGAG 247001
```

QY	2018	AATTATTATG--TCTGTATAGATAAAGATTAAATATATATGTTAAAAAA	2065
Db	245940	AATTATTATGCTCTGTTACAGATAAAGATTATATATGTTAGTAA	245891
RESULT 15			
AC142108.2			
WFOCOMMENT			
Sequence split into 4 fragments LOCUS AC142108 Accession AC142108			
Fragment Name Begin End			
AC142108.0 1 110000			
AC142108.1 100001 210000			
AC142108.2 200001 310000			
AC142108.3 300001 377129			
Continuation (3 of 4) of AC142108 from base 200001 (AC142108 Mus musculus chromosome UNK			
Query Match 28.8%; Score 594.6; DB 2; Length 110000;			
Best Local Similarity 73.1%; Pred. No. 1.4e-94;			
Matches 986; Conservative 0; Mismatches 304; Indels 59; Gaps 15;			
QY	737	CTATCCAGGTTAACTTGATGTTAAATGGAACCCAGTAATATCTGCTTTGGGGGACA	796
Db	51523	CTTTCAGGTTAACTTGATGTTAAATGGAACCCAGTCATCTGCAATTTGCTGGTATA	51582
QY	797	AAAGATCTTACTGCTGAAGCTGCCAGATGGTCTGCTCTACCTAGATATAAGAGATA	856
Db	51583	AGAATGTGACCGTGAAGCGCCACTAATGGAGTCTCTGCTACCTGGACAAAGATA	51642
QY	857	AGGTTTACCTAAACTGGAGAAAGGTAATTTGGTGGAGGCTGGAGTATTCACGTTTT	916
Db	51643	AGGCTCTACCTAAACTGGAGAAAGGTAACCTGCTGGTGGCTGGCAGTATTCACGTTTT	51702
QY	917	CTGGCTTTCTGGTGTTCCTCTATAGAACCCGACCTCTCCGATGCCCAAGAGTGAG	976
Db	51703	CTGGCTTTCTGGTGTTCCTCTATAGAACCCGACCTCTCCGATGCCCAAGAGTGAGG	51762
QY	977	GGATGCCACCTCTCGATGTTATGGAAGATCAATTTTTCATCATTTGGATGATGCTTTT	1036
Db	51763	CG-----CCACCCCGGTTATCAGAGAGCACTCTTGGATGATGCTCTTACTGTTTT	51817
QY	1037	ATTGGTTTCTATGGGTTGATGATTTCTAAGATTCAG--CTGTCTGGAACCAATACA	1095
Db	51818	GTTTCATGGC--TGACTATGGATTCCTCTGAGAATCTAGACCTCTCTGGACAGACACA	51874
QY	1096	AAATTTTACAGATTATTT---TGTGTGTCTCTTTTCAATATTTTGGATGGGACTCTA	1151
Db	51875	AGTTTTCACAGATCAATTTCTGTGTGTGTCAGGTTTCTTATATGTTGGAGCTGGACTCTA	51934
QY	1152	AGCAGATAATACC-----TATGCTTAAATGTAAACAGTCAAAAGCTGTCTGCAAGACTTAT	1206
Db	51935	AAAGCTGAAATAATCCCGATCTGCTTACATGTTAACTGTCAAGAGCCGCTCTCAAGGTTGT	51994
QY	1207	TCGGAATTCATTTCTGGGATTTAGTAATTTAGTACAGATGTGGAA---TTTTATTTGT	1263
Db	51995	CCGTG--TTTAGTTTCTGCTGTTTACAGCATTTGGCTGCAGATGTGGGTTTTTTTATTTCT	52052
QY	1264	TTAGTTTAAAGACTGGCAACAGGCTCTAAGGATTTAGAAACTCTAAAGTTCTGACTTC	1323
Db	52053	TTGTTTTGAAGACAGGCTTAATGGGTTCTGAAGTTAGAAACTTGGAGTTCTGACTTC	52112
QY	1324	AATCAAGGTTAGTGTGATGACTGCCAAGAACTGTATAC-----TGTTG 1366	
Db	52113	AGT-----GGTATAGTGAACCTGCCAAGACATGATACAGTGTGGATACATGATTAT	52168
QY	1367	TTAATAATTGATTAATTTGTTTTTATTTCTTTGGAAATTTAGTTTGTGTTCTTGTGTA	1426
Db	52169	TAATTAATACITTTTAAACAAATTAATAATTTTCTTTTGAATTTGTTTGTCTCTTAA	52228
QY	1427	AAACTTGGATTTTTTTTTCAGTAACCTGATTTATGTTTCTTCTTAAATAAGGTAATG	1486
Db	52229	GGTAACCTGGGATTTGATTTTT--AGTGACTGGGTTGTGTTCTCTACCCATAAGGTAACA	52287
QY	1487	AATGGCTTGCCCAAAATTTTACCTTTGACTAGCATATCATCGACATGACTTCTCT-----CA	1542

Search completed: May 15, 2004, 09:06:12
Job time : 8117 secs

Db	52288	AATGCTTTGGCTCAAAATGTCCTTTAACTCTGAACCTCACCACATGACCTCCCTCCTGCC	52347
QY	1543	AAAAAAAAGATGCTTTCATAGTTGATTTTAAATGATATATGGAAGAGTCAATATTTCC	1602
Db	52348	CCTAAACGAATGCTCCATAGTTGATTTTAAATATATATATGTAAGAGTCAATATTTCC	52407
QY	1603	AAGTTATATTTTCTAAGAAAGAAATAGATCATATAATCTGCAAGGAAAAAGTTGCTTAC	1662
Db	52408	AAATTTATATTTTCTAATAGGAAAAACAGATCAT--NATCTGCAAGAAAGAAAGTTATTTAC	52466
QY	1663	CCAAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAAACAGCTCCCTCCGAGGGAATC	1722
Db	52467	CCCAACTCTAA--TGCTCAGCGCCAAACCTTGGCAAAAGTACCTCTCTTGTGGGAAATC	52524
QY	1723	TTATATCTTTTATGCTCAACTTTTAAATAAATGATTGATAATAACCACTTTATTAATAAAC	1782
Db	52525	GAATACTCTATCGCTCAACCTTTAAATAACATGATTGATAATAGCTACTTTTATTAACCA	52584
QY	1783	TAAAGTTTTTTTTTTTTTTTCCTAGACATGACCACTTTTATTAACCTGGTGGATGCTGTT	1842
Db	52585	TAAAG--AATATTTTCTCTTCACTATGACCACTTTTAACTGGAGATAGTATGCCCTT	52642
QY	1843	GTTTCTAATTATACCTAATTTTTCAAGGCTTCTGTTGTAATTTGAAGTATCATCTGGTTTG	1902
Db	52643	ATTTTTAATTGTATCTAATTTTTCAAACTCTCTGTTGTTGTTGAAATATCATCTGGTTTG	52702
QY	1903	CCTTAACCTTTTAAATTTGTATATATTTATCTGTTTGTAGCTAATATAAATTCAAATATCC	1962
Db	52703	CCTTAATTCCTTAACTTGTATATATTTATCTGTTCAAGCTAATATAAATTCAGATACTCC	52762
QY	1963	ATATCTAAATTTAGTGCAATATC-----TTGTCTTTTGTATAGGTCATATGAATTCATAA	2017
Db	52763	CTAACTAAATTTAGTGCAATATCTTATTTTGTCTTTTGTATAGGTCATATGAATTCATAA	52822
QY	2018	AAATTTATGCTGTTTATAGAAATAAAGA 2046	
Db	52823	AAATTTATGCTGCTCTGTTTAAACAAATAAA 52851	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 15, 2004, 04:14:15 ; Search time 5398 Seconds
(without alignments)
11423.750 Million cell updates/sec

Title: US-10-015-387A-219

Perfect score: 2065

Sequence: 1 gfgaagtggaggtttgatg.....attaatatgttaaaaaa 2065

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2065	100.0	3091	11	BC050026 Homo sapi
2	1077	52.2	2406	11	AK032406 Mus muscu
3	1012.8	49.0	2802	11	AK032621 Mus muscu
4	751	36.4	905	12	BI668562

5	686	33.2	696	13	EX093493
6	606	29.3	606	29	AY418431
7	603	28.2	606	29	AY418432
8	592.6	28.7	706	9	AV722278
9	572.4	27.7	579	13	EX506987
10	520.4	25.2	700	12	BI669085
11	504	24.4	504	9	AA868507
12	474	23.0	597	29	AY418433
13	454.8	22.0	745	14	CB318208
14	399.8	19.4	436	13	BU781883
15	396	19.2	750	13	BU382466
16	388.4	18.8	734	13	BU369123
17	388	18.8	421	9	AV721267
18	375.2	18.2	683	9	AV343573
19	357	17.3	597	9	AJ452689
20	334	16.2	340	9	AL732165
21	326.2	15.8	331	9	AI468259
22	323.6	15.7	566	10	BE755791
23	320.8	15.5	324	9	AW074518
24	315.6	15.3	674	9	AV348597
25	311.8	15.1	802	13	BQ444664
26	307.4	14.9	756	14	CE207272
27	292.4	14.2	1783	11	AK042555
28	292.4	14.2	2202	11	AK032112
29	290.4	14.1	754	13	BU140889
30	288	13.9	903	13	BO722274
31	287.6	13.9	346	9	AW047113
32	280	13.6	291	10	BF445600
33	275.4	13.3	689	13	BU457636
34	262.8	12.7	284	9	AL449740
35	259.8	12.6	513	13	BY372910
36	258.8	12.5	447	9	AI848962
37	253.4	12.3	1185	13	EX394263
38	247.8	12.0	907	14	CD254575
39	244.4	11.8	440	10	BB311702
40	240.4	11.6	380	13	BY630619
41	239.2	11.6	717	12	BM950922
42	238.2	11.5	635	10	BB622353
43	237.6	11.5	883	14	CD254586
44	235.4	11.4	422	10	BE779670
45	233.6	11.3	2428	11	AK076174

ALIGNMENTS

RESULT 1
BC050026

LOCUS

DEFINITION

Homo sapiens,

3091 bp

mRNA

linear

HTC 10-APR-2003

clone IMAGE:4820943,

BC050026

BC050026

BC050026.1

GI:29747708

HTC.

KEYWORDS

SOURCE

Homo sapiens

Homo sapiens

Homo sapiens

ORGANISM

REFERENCE

1 (bases 1 to 3091)

Strausberg,R.

Direct Submission

Submitted (01-APR-2003)

National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MSC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 71 Row: c Column: 9
 This clone has the following problem: retained intron.

FEATURES

source
 Location/Qualifiers
 1..3091
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4820943"
 /tissue_type="Brain, hippocampus"
 /clone_lib="NIH_MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Query Match 100.0%; Score 2065; DB 11; Length 3091;

Best local similarity 100.0%; Pred. No. 0;

Matches 2065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1  GTCAATGTAGGTTTGATCACTTCAGATCTCTAGAACACAGATGGGTGCGAGGGCCC 60
Db  GTGAAATGTAGGTTTGATCACTTCAGATCTCTAGAACACAGATGGGTGCGAGGGCCC 1027

QY 61  CAGCGAGGGCTGATCTTGGCGGAGAGAGTAGGTAAAGGGTCTCGATGAGTCCTT 120
Db  CAGCGAGGGCTGATCTTGGCGGAGAGAGTAGGTAAAGGGTCTCGATGAGTCCTT 1087

QY 121 AAGGACAAAGGTAAACAGACCGAGAGAGAGCTCGAGGGGAGACTTTGACTCAAGCCA 180
Db  AAGGACAAAGGTAAACAGACCGAGAGAGAGCTCGAGGGGAGACTTTGACTCAAGCCA 1147

QY 181 CAGAATGTGGAGTGTGCGCGCGCGCGCGCGCGCTCGTCTCGACGCGTGTGACCTA 240
Db  CAGAATGTGGAGTGTGCGCGCGCGCGCGCGCGCTCGTCTCGACGCGTGTGACCTA 1207

QY 241 GCGCGTAGCATCTTCCGAGCACCGGATCCCGGGTAGGAGCGACGCGGGGAGCACC 300
Db  GCGCGTAGCATCTTCCGAGCACCGGATCCCGGGTAGGAGCGACGCGGGGAGCACC 1267

QY 301 AGCGCCAGCGGCTGCGGCTGCCACACGCGCTACCATGGGCTCCGGGCGCGGGCGCTG 360
Db  AGCGCCAGCGGCTGCGGCTGCCACACGCGCTACCATGGGCTCCGGGCGCGGGCGCTG 1327

QY 361 TCCGCGGTGCGGCGGTGCTGTGCTCTCACGCTGCGGGGCTGCCGCTGTGGGACAG 420
Db  TCCGCGGTGCGGCGGTGCTGTGCTCTCACGCTGCGGGGCTGCCGCTGTGGGACAG 1387

QY 421 AAGCACAGGAGGCCATCTGCTGTGGAGGCAAGTGTCTGTGTGTGCTGAGTCTGAA 480
Db  AAGCACAGGAGGCCATCTGCTGTGGAGGCAAGTGTCTGTGTGTGCTGAGTCTGAA 1447

QY 481 GCCACGACTCCAGGAGCTCTCTCTCTCCCGCTGGGGATATCGGTCCGGGCGGCCAAC 540
Db  GCCACGACTCCAGGAGCTCTCTCTCTCCCGCTGGGGATATCGGTCCGGGCGGCCAAC 1507

QY 541 TCCAGGTGCGCTTCTCGGGGCTCGGAGCACCAACACGAGCCATCCGAGATGAGCAAC 600
Db  TCCAGGTGCGCTTCTCGGGGCTCGGAGCACCAACACGAGCCATCCGAGATGAGCAAC 1567

QY 601 AAGACGGCGATCATTTACTTTCGATCAGATCTCTGTGAATGTGGGTAAATTTTTCACAT 660
Db  AAGACGGCGATCATTTACTTTCGATCAGATCTCTGTGAATGTGGGTAAATTTTTCACAT 1568

QY 661 GAGTCTGCTTTGTAGCACCAAGAAAGGAATTTACAGTTTCAGTTTTCAGTGATTA 720

```

```

Db 1628 GAGTCTGCTTTGTAGCACCAAGAAAGGAATTTACAGTTTCAGTTTTCAGTGATTA 1687
QY 721 GTCTACAGAGCAAACTATCCAGGTAACTTGATGTTAAATGGAAGAAACAGTAATCT 780
Db 1688 GTCTACAGAGCAAACTATCCAGGTAACTTGATGTTAAATGGAAGAAACAGTAATCT 1747
QY 781 GCCTTTGCGGGGACAAAGATGTTACTCGTGAAGCTGCCACGAATGCTCTCTGCTCTAC 840
Db 1748 GCCTTTGCGGGGACAAAGATGTTACTCGTGAAGCTGCCACGAATGCTCTCTGCTCTAC 1807
QY 841 CTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGGTAATTTGGTGGAGCGTGG 900
Db 1808 CTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAGAAAGGTAATTTGGTGGAGCGTGG 1867
QY 901 CAGTATTCACGTTTCTGGCTTTCTGGTGTTCCTCCCTATAGGATTCATTTCTCCATGA 960
Db 1868 CAGTATTCACGTTTCTGGCTTTCTGGTGTTCCTCCCTATAGGATTCATTTCTCCATGA 1927
QY 961 TGTTCATCCAGGTGAGGGATGACCCACTCTCAGTGTATTGGAGAGATCATTTTTCATCAT 1020
Db 1928 TGTTCATCCAGGTGAGGGATGACCCACTCTCAGTGTATTGGAGAGATCATTTTTCATCAT 1987
QY 1021 TGGATTGATGTCCTTTTATTGGTTTCTCATGGTGGATATGGATTCTAAGGATTTCTAGCCT 1080
Db 1988 TGGATTGATGTCCTTTTATTGGTTTCTCATGGTGGATATGGATTCTAAGGATTTCTAGCCT 2047
QY 1081 GTCTGAACCAATACAAAATTCACAGATTATTGTGTGTGTCTGTTTCAGTATATTGGA 1140
Db 2048 GTCTGAACCAATACAAAATTCACAGATTATTGTGTGTGTCTGTTTCAGTATATTGGA 2107
QY 1141 TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAAGCTGTCTCAAG 1200
Db 2108 TTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTAACAGTCAAAAGCTGTCTCAAG 2167
QY 1201 ACTTATCTGAATTCATTTCTGGGATTACTGAATTAAGTTCAGATGTGGAATTTTAT 1260
Db 2168 ACTTATCTGAATTCATTTCTGGGATTACTGAATTAAGTTCAGATGTGGAATTTTAT 2227
QY 1261 TGTTTAGTTTAAAGACATGGCAACCCAGCTCTAAGGATAGAAAACCTCTAAGTCTGAC 1320
Db 2228 TGTTTAGTTTAAAGACATGGCAACCCAGCTCTAAGGATAGAAAACCTCTAAGTCTGAC 2287
QY 1321 TTCAATCAACGGTGTAGTGATACCTGCCAAAGACTGTATACCTGTGTTAATATTGATT 1380
Db 2288 TTCAATCAACGGTGTAGTGATACCTGCCAAAGACTGTATACCTGTGTTAATATTGATT 2347
QY 1381 ATATTTGTTTATTCCTTTTGGAAATTAAGTGTGTTGTTTCTGTGTAAGAACTTGGATT 1440
Db 2348 ATATTTGTTTATTCCTTTTGGAAATTAAGTGTGTTGTTTCTGTGTAAGAACTTGGATT 2407
QY 1441 TTTTTCAGTAACTGGTATTATGTTTCTCTTAAATAAGGTAATGAATGGCTTGCCAC 1500
Db 2408 TTTTTCAGTAACTGGTATTATGTTTCTCTTAAATAAGGTAATGAATGGCTTGCCAC 2467
QY 1501 AAATTTACCTTGACTACGATATCATCGACATGCTCTCTCAAAAAGAAAGATGTTCA 1560
Db 2468 AAATTTACCTTGACTACGATATCATCGACATGCTCTCTCAAAAAGAAAGATGTTCA 2527
QY 1561 TAGTGTGTTTAAATTTGATATGTGAAGAGTCATATTTTCCAGTTATATTCTTAAGA 1620
Db 2528 TAGTGTGTTTAAATTTGATATGTGAAGAGTCATATTTTCCAGTTATATTCTTAAGA 2587
QY 1621 AGAAGAAATAGATCATAAATCTGACAGGAAAGGTTGCTTACCAGAAATCTAAGTCTCA 1680
Db 2588 AGAAGAAATAGATCATAAATCTGACAGGAAAGGTTGCTTACCAGAAATCTAAGTCTCA 2647
QY 1681 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAATCTTATATCTTTATGCTCAA 1740
Db 2648 ATCCCTGAGCCTCAGCAAAACAGCTCCCTCCGAGGGAATCTTATATCTTTATGCTCAA 2707
QY 1741 CTTTAAATTAATGATTGATTAATAACACCTTTATTAAGAACTTAAGGTTTTTTTTTTC 1800

```

Db 2708 CTTTAATAAATGATGATAATAAACCACCTTTTATTAATAAACCTTAAGGTTTTTTTTTTTC 2767

QY 1801 CGTAGACATGACACCTTTTATTAACCTGCTGGGATGCTGTGTTCTTAATATATACCTAT 1860

Db 2768 CGTAGACATGACACCTTTTATTAACCTGCTGGGATGCTGTGTTCTTAATATATACCTAT 2827

QY 1861 TTTTCAAGGCTTCTGTTGATTTTGAAGTATCATCTGCTGTTTGGCTTAACTCTTTAAATG 1920

Db 2828 TTTTCAAGGCTTCTGTTGATTTTGAAGTATCATCTGCTGTTTGGCTTAACTCTTTAAATG 2887

QY 1921 TATATATTTATCTGTTAGCTAATATTAATTCAAATATCCCATCTCTAAATTTAGTGCA 1980

Db 2888 TATATATTTATCTGTTAGCTAATATTAATTCAAATATCCCATCTCTAAATTTAGTGCA 2947

QY 1981 ATATCTGTCTTTGTATAGTCAATGAATTCATAAAATTTATTAATGCTGTATAGAA 2040

Db 2948 ATATCTGTCTTTGTATAGTCAATGAATTCATAAAATTTATTAATGCTGTATAGAA 3007

QY 2041 TAAAGATTAATATGTTAAAAAA 2065

Db 3008 TAAAGATTAATATGTTAAAAAA 3032

RESULT 2

AK032406

LOCUS

DEFINITION

AK032406 2406 bp mRNA linear HTC 18-SEP-2003

Mus musculus adult male olfactory brain cDNA, RIKEN full-length

ENRICHED LIBRARY, clone:6430538G21 product:CEREBELLIN-LIKE

GLYCOPROTEIN PRECURSOR, full insert sequence.

ACCESSION

AK032406

VERSION

AK032406.1 GI:26328228

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PubMed

10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

MEDLINE

20499374

PubMed

11042159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

MEDLINE

20530913

PubMed

11076961

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

6 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

7 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

8 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

9 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

10 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

11 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

12 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

13 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

14 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

15 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

16 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

17 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

18 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

19 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

20 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

21 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

22 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

23 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

24 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

25 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

26 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

27 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

28 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

29 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

30 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

31 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

32 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

33 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

34 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

35 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

36 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

37 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

38 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

39 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

40 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

41 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

42 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

43 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

44 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

45 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

46 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

47 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

48 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

49 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

50 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

51 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

52 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

53 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

54 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

55 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

56 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

57 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

58 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

59 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

60 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

61 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

62 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

63 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

64 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

65 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

66 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

67 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

68 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

69 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

70 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

71 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

72 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

73 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

74 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

75 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

76 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

77 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

78 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

79 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

80 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

81 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

82 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

83 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

84 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

85 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

86 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

87 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

88 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

89 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

90 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

91 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

92 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

93 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

94 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

95 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

96 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

97 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

98 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

99 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

100 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

101 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

102 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

103 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

104 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

105 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

106 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

107 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

108 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

109 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

110 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

111 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

112 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

113 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

114 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

115 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

116 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

117 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

118 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

119 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

120 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

121 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

122 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

123 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

124 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

125 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

126 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

127 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

128 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

129 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

130 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

131 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

132 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

133 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

134 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

135 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

136 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

137 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

138 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

139 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

140 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

141 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

142 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

143 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

144 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

145 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

146 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

147 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

148 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

149 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

150 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

151 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

152 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

153 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

154 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

155 Nature 420, 563-573 (2002)

JOURNAL

MEDLINE

11076961

PubMed

REFERENCE

156 Nature 420, 563-573 (2002)

JOURNAL

SOURCE ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	Carninci, P. and Hayashizaki, Y.
JOURNAL	High-efficiency full-length cDNA cloning
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)
PUBMED	99279253
REFERENCE	10349636
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K., Yoneda, I., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, M., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
PUBMED	11076861
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
PUBMED	12082166
REFERENCE	6 (bases 1 to 2802)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komori, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohmoto, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/. Location/Qualifiers
FEATURES	1. .2802

591 GATGAGCAACAGACGCGCATCATTTACTTCGATCAGATCCTGGTGAATGTGGTAAATTT 650
1417 GATGAGCAACAGACGCGCATCATTTACTTCGATCAGATCCTGGTGAATGTGGTAAATTT 1476
651 TTTCACATTTGGAGTCTGTCTTTGTAGACCAAGAAAGGAATTTACAGTTTCAGTTTCA 710
1477 TTTCACATTTGGAGTCTGTCTTTGTAGACCAAGAAAGGAATTTACAGTTTTCAGTTTCA 1536
711 CGTGATTTAAAGTCTACAGAGCCAACTATCCAGGTTTAACTGATGTTAAATGGAAC 770
1537 CGTAATTTAAAGTCTACAGAGCCAACTATCCAGGTTTAACTGATGTTAAATGGAAC 1596
771 AGTAATATCTGCTTTTCGGGGGACAAAGATGTTACTCGTGAAGCTGCCAGAAATGGTGT 830
1597 AGTCACTCTGCTGATTTGCTGTTGATAAGATGTAACCGGTGAAGCCGCCACTAATGGAGT 1656
831 CCGTCTCTACCTAGATAAGAGAGATAAGGTTTAACTTAAACTGAGAGAAAGGTAATTTGGT 890
1657 GCTCCTGTACCTGGACAAAGATAGGTTTAACTTAAACTGAGAGAAAGGTAATTTGGT 1716
891 TGGAGCTGCGCAGTATTCACAGTCTTTCTGGCTTTCTGGTCTCCCTATAGGATCAAT 950
1717 CGGTGCTGCGCAGTATTCACAGTCTTTCTGGCTTTCTGGTCTCCCTATAGAACCGGAC 1776
951 TTCTCCATGATGTTCAATCCAGGTGAGGATGACCCCACTCTCGATGTTATTGGAAGATCAAT 1010
1777 TTCTCCGATGATGCCAAGAGTGAAGGCG-----CCACCCCGGTTTATCAGAAGAGCACT 1831
1011 TTTCATCATGTAATGATGTCCTTTTATGTTTCTGTTTCTCATGTTGATGATCTTAAGS 1070
1832 CTTGGATGTAATGTCCTTACTGTTTGTTCATGCG---TGACTATGATTTCTCTCTGAGA 1888
1071 ATTCTAG-CTGTGCTGAACCAATACAAAATTTTCACAGATTAATTT---TGTTGTGTCTGT 1125
1889 ATTCTAGACTGTCTGACACAGACACAAAGTTTCACAGATCAATTTCTGTGTGTGCGCAGT 1948
1126 TTCAATATTTGGATGGAAGTCTGAGTCTAAGCAGATAATACC-----TATGCTTAATATGTAAC 1180
1949 TTCTTATATGTTGAGTGGGACTCTAAGCTGAATAATCCCGATCTGCTTACATGTTAA 2008
1181 AGTCAAGCTGTCTGCAAGACTATTTCTGAATTTCAATTTCTGTTGGAATTAATGAT 1240
2009 CTGTGAGAGCCGTCTCAAGTTTGTCTG----TTAGTTTCTGCTGTTACAGATGTCG 2066
1241 TACAGATGTTGGAAT---TTTATTTGTTTGTAGTTTAAAGACTGCGACACAGGCTCTAAGGA 1297
2067 TGCAGATGTTGCGTTTCTTTATTTCTTTGTTTGTAAAGACAGCTAATGCTGTAAG 2126
1298 TTGAATACTCTAAAGTTCTGACTTCAATCAACGTTAGTGTGATCTGCCAAGAACTG 1357
2127 GTAGAAACTTGAAGTTCTGACTTCAGT---GGTTATAGTGAACCTGCCAAGACATG 2182
1358 TATAC-----TGTTGTAATATATGATTAATTTCTGTTTATTTATCTTT 1400
2183 TATACAGTGTGATGATGATTAATTAATTAATTTTAAACAATTAATTAATTT 2242
1401 GGAATTAATTTGTTGTTCTCTGTAAGAACTGTTGTTTCTTCTGTTTCTGTTTCTGTTT 1460
2243 TGAATTTGTTGTTTCTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 2301
1461 ATGTTTCTCTTAAATTAAGTATGATGCTGTTGCCAACAATTAATTAATTAATTAATTT 1520
2302 GTGTTCTCTTACCCTAAGTATGATGCTGTTGCCAACAATTAATTAATTAATTAATTT 2361
1521 ATCATCGACATGATCTCTCAAAA-----AAAAAGATGCTTCTGATGTTGTTTAAATTT 1576
2362 CTCACCATGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2421
1577 GTATATGTAAGAGTATATTTTCCAGTTTATATTTTCTAAGAGAGAAATAGATCATA 1636
2422 ATATATGTAAGAGTATATTTTCCAAATTTATATTTTCTAATAGGAAAAACAGATCAT- 2480
1637 AATCTGACAGAAAGAGTTGCTTACCCCAAAATCTAAGTCTCAATCCCTGAGCCTCAGC 1696

2481 AATCTGACAAAGAGATTAATTTACCCCAACTCTAA---TGCTCAGCGCAAAACCTTGGC 2538
1697 AAAACAGCTCCCTCCGAGGAAATCTTATCTTTTCTCAACTTAAATTAATTAATGAT 1756
2539 AAGTACCTCTCTCTGGTGGAAATCGAATATCTATCTCTCAACTTAAATTAATGAT 2598
1757 TGATAATAACCACTTTATTTAAAAACCTAAGGTTTTTTTTTTTCCGTAGACATGACCCT 1816
2599 TGATAATAGCTACTTTATTAACCCATAAGG---AATATTTTCTCTTCAATGACCCT 2656
1817 TTATTAACCTGGTGGGATGCTGTTTCTTAAATTAATTAATTAATTAATTAATTAATTAAT 1876
2657 TTATTAACCTGGAGATGATGCTTATTTTAAATTTGATCTATTTTCAAAATCTCTCT 2716
1877 TGATTTGAAGTATCATCTGTTTGGCTTAACTCTTTAAATTTGATATATTTATCTGTT 1936
2717 TGTTTGAATATATCTGTTTGGCTTAACTCTTAACTTAACTTAACTTAACTTAACTTAACT 2776
1937 TAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1962
2777 CAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2802

RESULT 4
BI668562
LOCUS 603293849F1 NIH_MGC_96 905 bp mRNA linear EST 12-SEP-2001
DEFINITION mRNA sequence.
ACCESSION BI668562
VERSION BI668562.1 GI:15582782
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC <http://img.ncbi.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D. (NHGRI), Shiraki
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11793 row: b column: 14
High quality sequence stop: 735.

FEATURES
source
1..905
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5312941"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match		36.4%; Score 751; DB 12; Length 905;
Best Local Similarity		91.8%; Pred. No. 1.9e-124;
Matches 827; Conservative		0; Mismatches 70; Indels 4; Gaps 3;
QY	1055	GATATGGATTTCTAAGGATTTCTAGCTGTCTGACCAATACAAATTTTACAGATTTATTTG 1114
Db	4	GGTATGGATTTCTAAGGATTTCTAGCTGTCTGACCAATACAAATTTTACAGATTTATTTG 63
QY	1115	TGTGTCTGTCTGATGATTTTGGATTTGGGACTCTAAGCAGATTAATACCTATGCTTAAA 1174
Db	64	TGTGTCTGTCTGATGATTTTGGATTTGGGACTCTAAGCAGATTAATACCTATGCTTAAA 123
QY	1175	TCTAACAGTCAAAAGCTGTCTCAAGACTTATTTCTGAATTTTCCTGGGATTTACTGA 1234
Db	124	TCTAACAGTCAAAAGCTGTCTCAAGACTTATTTCTGAATTTTCCTGGGATTTACTGA 183
QY	1235	ATTAGTTACAGATGTGGAAATTTTATTTGTTAGTTTAAAGACTGGCAACAGGCTCTAA 1294
Db	184	ATTAGTTACAGATGTGGAAATTTTATTTGTTAGTTTAAAGACTGGCAACAGGCTCTAA 243
QY	1295	GGATTAGAAAACCTCTAAAGTCTGACTTCAATCAACGGTTAGTGTGATATGCCAAGAA 1354
Db	244	GGATTAGAAAACCTCTAAAGTCTGACTTCAATCAACGGTTAGTGTGATATGCCAAGAA 303
QY	1355	CTGTATACGTGTTAATATATGATATATTTGTTTATTTCTCTTGGAAATTTAGTTTGT 1414
Db	304	CTGTATACGTGTTAATATATGATATATTTGTTTATTTCTCTTGGAAATTTAGTTTGT 363
QY	1415	TGGTCTCTGTAAGAACTTGGATTTTCTTTCAGTAACTGGTATTTATGTTTCTCTTAA 1474
Db	364	TGGTCTCTGTAAGAACTTGGATTTTCTTTCAGTAACTGGTATTTATGTTTCTCTTAA 423
QY	1475	AATAAGGTAATGAATGCGTTGCCCAAAATTTACCTTGTACTAGATATCATCGACATGAC 1534
Db	424	AATAAGGTAATGAATGCGTTGCCCAAAATTTACCTTGTACTAGATATCATCGACATGAC 483
QY	1535	TTCTCTCAAAAAAAGAAATGCTTCATAGTTGTATTTTAAATGTATATGAAAGAGTCA 1594
Db	484	TTCTCTCAAAAAAAGAAATGCTTCATAGTTGTATTTTAAATGTATATGAAAGAGTCA 543
QY	1595	TATTTTCCAAAGTTATTTTCTAAGAAAGAAATGATCATTAATCTGACAGGAAAGAG 1654
Db	544	TATTTTCCAAAGTTATTTTCTAAGAAAGAAATGATCATTAATCTGACAGGAAAGAG 603
QY	1655	TTGCTTACCCAAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAACAGCTCCCTCCGA 1714
Db	604	TTGCTTACCCAAAATCTAAGTGTCTCAATCCCTGAGCTCAGCAAAACAGCTCCCTCCGA 563
QY	1715	GGAAATCTTATACCTTTATTTGCTCAACTTTAATTAATAATGATT--GATAATAACCACTTT 1772
Db	664	GGAAATCTTATACCTTTATTTGCTCAACTTTAATTAATAATGATTAAACCCACTTT 723
QY	1773	ATTAAAACTTAAGTTTTTTTTTTTTCGTAGACATGACCACCTTTATTAACCTGTGGTG 1832
Db	724	TATTTAAAACTTAAGGGTGTTTTTTCCCGTAAACCTTGACCCCAATTTATTAACCTGGCGGG 783
QY	1833	GGATGCTGTGTCTCAATATATACCTTATTTTCAAGCTTCTGTCTGTTTGAAGTATCA 1892
Db	784	GGATGCTGTGTCTCAATATATACCTTATTTTCAAGCTTCTGTCTGTTTGAAGTATCA 842
QY	1893	CTGG--TTTTGCTTAACTCTTTAAATTTGATATATTTATCTGTTTGTAGCTTAATTAAT 1951
Db	843	CCGGCTTTGGCCCTTAACTCTTTAATTTGATATTTTAAATCCGGTGAATTTAAATTTCCAT 902
QY	1952	T 1952
Db	903	T 903
RESULT 5		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		
DEFINITION		
EST 23-JAN-2003		
BX093493		
LOCUS		

```
Db 241 AGAAGAAGATAGATCATTAATCTGACAGGAAAAAGTTGCTTACCCAAAACTAAGTGC 300
QY 1678 TCAATCCCTGAGCTCAGCAAAAAGCTCCCTCCGAGGGAATCTTATCTTATTTGCT 1737
Db 301 TCAATCCCTGAGCTCAGCAAAAAGCTCCCTCCGAGGGAANCTTATCTTATTTGCT 360
QY 1738 CAATCTTAATAAATGATGATAAATAACACCTTTATTAATAAACCCTAAGGTTTTTTT 1797
Db 361 CAATCTTAATAAATGATGATAAATAACACCTTTATTAATAAACCCTAAGGTTTTTTT 420
QY 1798 TTCGCTAGACATGACACCTTTATTAATCTGCTGGTGGATGCTGCTGCTTCTAATATACC 1857
Db 421 TTCGCTAGACATGACACCTTTATTAATCTGCTGGTGGATGCTGCTGCTTCTAATATACC 480
QY 1858 TATTTTCAAGGCTCTGTTGTTATTTGAAGTATCATCTGCTTTTGCTTAACTCTTTAAA 1917
Db 481 TATTTTCAAGGCTCTGTTGTTATTTGAAGTATCATCTGCTTTTGCTTAACTCTTTAAA 540
QY 1918 TTGTATATATTTATCTGTTTAGCTATATTAATAATCCATATCCCATATCTAATTTAGT 1977
Db 541 TTGTATATATTTATCTGTTTAGCTATATTAATAATCCATATCCCATATCTAATTTAGT 600
QY 1978 GCAATATCTGCTCTTTGTTAGTCTCATATGAAATTCATAAAATTTATGCTGTTATA 2037
Db 601 GCAATATCTGCTCTTTGTTAGTCTCATATGAAATTCATAAAATTTATGCTGTTATA 660
QY 2038 GAATAAGATTAATATATATGTTAAAAAAA 2065
Db 661 GAATAAGATTAATATATGTTAAAAAAA 688
```

```
RESULT 6
AY418431
LOCUS
DEFINITION
Homo sapiens CBLN1 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION
AY418431
VERSION
AY418431.1 GI:39774391
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 606)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED
14671302
REFERENCE
2 (bases 1 to 606)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submision
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..606
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>606
/gene="CBLN1"
/locus_tag="HCM655"
gene
29.3%; Score 606; DB 29; Length 606;
ORIGIN
Query Match
```

```
Best Local Similarity 100.0%; Pred. No. 1.9e-98;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 337 ATGGGCTCCGGGCGCGGGCGCTGTCGGGGTCCGGCGCTGCTGCTCAGCGTG 396
Db 1 ATGGGCTCCGGGCGCGGGCGCTGTCGGGGTCCGGCGCTGCTGCTCAGCGTG 60
QY 397 CCGGGCTCCCGCTCTGGGCAACAGACACGAGGCCCATCGTGTGAGAGGCAAGTGT 456
Db 61 CCGGGCTCCCGCTCTGGGCAACAGACACGAGGCCCATCGTGTGAGAGGCAAGTGT 120
QY 457 CTGGTGTGTGGACATCGAACCCGGCCACGAGCTCCAGGGTCTCTTCTCCCGCTG 516
Db 121 CTGGTGTGTGGACATCGAACCCGGCCACGAGCTCCAGGGTCTCTTCTCCCGCTG 180
QY 517 GGGATATCGGTCCGGGCGGCCAATCCCAAGTCTGCTTCTCGGGTGGGAGCAAC 576
Db 181 GGGATATCGGTCCGGGCGGCCAATCCCAAGTCTGCTTCTCGGGTGGGAGCAAC 240
QY 577 CACGAGCCATCCGAGATGAGCAACAGACGCGCATCATTTACTTCGATCAGATCCTGGTG 636
Db 241 CACGAGCCATCCGAGATGAGCAACAGACGCGCATCATTTACTTCGATCAGATCCTGGTG 300
QY 637 AATGGGGTAATTTTTCATGAGTCTGCTTTGTAGACACCAAGAAAGAAATTTAC 696
Db 301 AATGGGGTAATTTTTCATGAGTCTGCTTTGTAGACACCAAGAAAGAAATTTAC 360
QY 697 AGTTTCAGTCTTTCAGTGAATTAAGTCTACAGAGCCAACTATCCAGGTAACTTGATG 756
Db 361 AGTTTCAGTCTTTCAGTGAATTAAGTCTACAGAGCCAACTATCCAGGTAACTTGATG 420
QY 757 TTAATGGAACCAACAGTAAATATCTGCTTTGGGGGACAAAGATGTTACTCGTGAAGCT 816
Db 421 TTAATGGAACCAACAGTAAATATCTGCTTTGGGGGACAAAGATGTTACTCGTGAAGCT 480
QY 817 GCCAGAAATGGTGTCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAG 876
Db 481 GCCAGAAATGGTGTCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAAACCTGGAG 540
QY 877 AAAGGTAATTTGGTGGAGGCTGGCAGTATCCAGCTTTCTGGCTTTCTGGTGTTCGCC 936
Db 541 AAAGGTAATTTGGTGGAGGCTGGCAGTATCCAGCTTTCTGGCTTTCTGGTGTTCGCC 600
QY 937 CTATAG 942
Db 601 CTATAG 606
RESULT 7
AY418432
LOCUS
DEFINITION
Pan troglodytes CBLN1 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION
AY418432
VERSION
AY418432.1 GI:39774392
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 (bases 1 to 606)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED
14671302
REFERENCE
2 (bases 1 to 606)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
```



```

541 TAAGGATTCAGCTGCTGAACC-----ATACTCACAGATATTATGGGNGTGTCTGT 594
1126 TTCAGTATATTTGGATTGGGACTCTAGCAGATAAATACCTATCTTAATGTA 1178
595 TTCAGTATATTTGGATTGGGACTTTAAG--AGATAATCCTATGCTTAATGTA 645

RESULT 9
BX506987          579 bp      mRNA      linear      EST 04-SEP-2003
LOCUS             DKFZp799P072_r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION        BX506987          579 bp      mRNA      linear      EST 04-SEP-2003
ACCESSION         BX506987          579 bp      mRNA      linear      EST 04-SEP-2003
VERSION           BX506987.1 GI:32047784
KEYWORDS           EST.
SOURCE            Homo sapiens (human)
ORGANISM           Homo sapiens
REFERENCE          1 (bases 1 to 579)
AUTHORS            Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Wewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE             EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
JOURNAL            Unpublished (2003)
COMMENT            Contact: MIPS
MIPS              Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
Human Genome Project.
No si sequence available.
This clone (DKFZp799P072) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp799P072"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/notes="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN
Query Match          27.7%; Score 572.4; DB 13; Length 579;
Best Local Similarity 99.7%; Pred. No. 28-92;
Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 377 TGCTGTGTTGCTCTCACGCTGCGGGGTGCGCGTCTGGGCACAGACGACGAGCCCA 436
DB 1 TGCTGTGTTGCTCTCACGCTGCGGGGTGCGCGTCTGGGCACAGACGACGAGCCCA 60
QY 437 TCCTGTGTTGAGGGCAAGTGTCTGGTGTGTGCGACTCGAACCCCGGCCACGACTCCAAGG 496
DB 61 TCCTGTGTTGAGGGCAAGTGTCTGGTGTGTGCGACTCGAACCCCGGCCACGACTCCAAGG 120
QY 497 GCTCTCTCTCTCCCGCTGGGATATCGTTCGGGGGGCGCACTCAAGTGCCTTCT 556
DB 121 GCTCTCTCTCTCCCGCTGGGATATCGTTCGGGGGGCGCACTCAAGTGCCTTCT 180
QY 557 CGGCGGTGCGGAGCACCACACGAGCCATCCGAGATGAGCAACAGACGCGCATCATTT 616
DB 181 CGGCGGTGCGGAGCACCACACGAGCCATCCGAGATGAGCAACAGACGCGCATCATTT 240

```

```

QY 617 ACTTCGATCAGATCCCTGGTGAATGTGGTAAATTTTTCACATTGGAGTCTGTCTTTGTAG 676
DB 241 ACTTCGATCAGATCCCTGGTGAATGTGGTAAATTTTTCACATTGGAGTCTGTCTTTGTAG 300
QY 677 CACCAAGAAAAGGAATTTACAGTTTCAGTTTTCACGTGATTAAAGTCTACGAGGCCAA 736
DB 301 CACCAAGAAAAGGAATTTACAGTTTCAGTTTTCACGTGATTAAAGTCTACGAGGCCAA 360
QY 737 CTATCCAGGTTAACCTTGATGTTAAATGGAAACCAGTATATCTGCCTTTGGGGGGGACA 796
DB 361 CTATCCAGGTTAACCTTGATGTTAAATGGAAACCAGTATATCTGCCTTTGGGGGGGACA 420
QY 797 AAGATCTTACTCGTGAAGCTGCCACGAATGGTGTCTCTCTACCTAGATAAAGAGGATA 856
DB 421 AAGATCTTACTCGTGAAGCTGCCACGAATGGTGTCTCTCTACCTAGATAAAGAGGATA 480
QY 857 AGGTTTACCTAAACCTGGAGAAGGTAATTTGGTGGAGGCTGGCAGTATTCACGTTTTT 916
DB 481 AGGTTTACCTAAACCTGGAGAAGGTAATTTGGTGGAGGCTGGCAGTATTCACGTTTTT 540
QY 917 CTGGCTTCTCGTGTGTTCCCTTATAGGATTCATTT 951
DB 541 CTGGCTTCTCGTGTGTTCCCTTATAGGATTCATTT 575

RESULT 10
BX506985
LOCUS             603295195F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314390 5',
DEFINITION        mRNA sequence.
ACCESSION         BI669085
VERSION           BI669085.1 GI:15583318
KEYWORDS           EST.
SOURCE            Homo sapiens (human)
ORGANISM           Homo sapiens
REFERENCE          1 (bases 1 to 700)
AUTHORS            NIH-MGC http://imgc.nci.nih.gov/
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL            Unpublished (1999)
COMMENT            Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11796 row: n column: 23
High quality sequence stop: 581.
Location/Qualifiers
1..700
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5314390"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH MGC 96"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(Stcag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN

```



```

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 431.
Location/Qualifiers
1. .504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1408691"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/note="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5]
TGTTACCAATCTGAAGTGGAGCGCGCCAAATTTTTTTTTTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cots, and was
constructed by Renzo Soares and M. Fatima Bonaldo."

```

ORIGIN

Query Match	24.4%;	Score 504;	DB 9;	Length 504;	
Best Local Similarity	100.0%;	Pred. No. 3.8e-80;			
Matches 504;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1561	TAGTTGTATTTTAAATGTTATATGGAAGAGTCATATTTTCCAAAGTTATATTTTCTAAGA	1620		
DB	504	TAGTTGTATTTTAAATGTTATATGGAAGAGTCATATTTTCCAAAGTTATATTTTCTAAGA	445		
QY	1621	AGAGAATAGATCATATAATCTGCACAGGAAAAAGTTGCTTACCCAAAACTAAGTGCTCA	1680		
DB	444	AGAGAATAGATCATATAATCTGCACAGGAAAAAGTTGCTTACCCAAAACTAAGTGCTCA	385		
QY	1681	ATCCCTGAGCCTCAGCAAAAACAGCTCCCTCCGAGGGAAATCTTATACTTTTATTGCTCAA	1740		
DB	384	ATCCCTGAGCCTCAGCAAAAACAGCTCCCTCCGAGGGAAATCTTATACTTTTATTGCTCAA	325		
QY	1741	CTTTAAATTTAAATGATTGATTAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTTTC	1800		
DB	324	CTTTAAATTTAAATGATTGATTAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTTTC	265		
QY	1801	CGTAGACATGACCACTTTTATTAACTGGTGGGATGCTGTGTTTCTAAATATACCTAT	1860		
DB	264	CGTAGACATGACCACTTTTATTAACTGGTGGGATGCTGTGTTTCTAAATATACCTAT	205		
QY	1861	TTTTCAAGGCTCTGTGTATTTGAAGTATCATCTGGTTTTCCTTAAGTTTCTTTAAATG	1920		
DB	204	TTTTCAAGGCTCTGTGTATTTGAAGTATCATCTGGTTTTCCTTAAGTTTCTTTAAATG	145		
QY	1921	TATATATTATCTGTTTAACTAAATTAATTAATCAAAATCCCAATCTCAAAATTTAGTGCA	1980		
DB	144	TATATATTATCTGTTTAACTAAATTAATTAATCAAAATCCCAATCTCAAAATTTAGTGCA	85		
QY	1981	ATATCTTCTCTTTTGTATAGTGCATATGAATTCATAAAATTTATTTATGCTGTTATAGAA	204		
DB	84	ATATCTTCTCTTTTGTATAGTGCATATGAATTCATAAAATTTATTTATGCTGTTATAGAA	25		
QY	2041	TAAAGATTAAATATATGTTAAAAAA	2064		
DB	24	TAAAGATTAAATATATGTTAAAAAA	1		

RESULT 12	ACCESSION
AY418433	VERSION
LOCUS	KEYWORDS
DEFINITION	

Query Match	25.2%;	Score 520.4;	DB 12;	Length 700;
Best Local Similarity	96.1%;	Pred. No. 4.3e-84;		
Matches 544;	Conservative 0;	Mismatches 21;	Indels 1;	Gaps 1;
1168	GCTTAAATGTAACAGTCAAAAGCTGCTGCGAAGACTTATCTGAATTTCAATTTCCCTGGGA	1227		
5	GCTTAAATGTAACAGTCAAAAGCTGCTGCGAAGACTTATCTGAATTTCAATTTCCCTGGGA	64		
1228	TTACTGAAATTAGTTACAGATGCGAATTTTATTTGTTTAGTPTTAAAGACTCGCAACCA	1287		
65	TTACTGAAATTAGTTACAGATGCGAATTTTATTTGTTTAGTPTTAAAGACTCGCAACCA	124		
1288	GGCTTAAGGATTAGAAAACCTCTAAAGTCTCGACTCAATCAACGGTAGTGTGTAICTGC	1347		
125	GGCTTAAGGATTAGAAAACCTCTAAAGTCTCGACTCAATCAACGGTAGTGTGTAICTGC	184		
1348	CAAGAACCTGTATACCTGTGTTTAAATATATGATATATTTGTTTATTCCTTTGGGAATTA	1407		
185	CAAGAACCTGTATACCTGTGTTTAAATATATGATATATTTGTTTATTCCTTTGGGAATTA	243		
1408	GTTTGTTCGGTCTCTGTAAAAAACCCTGGATTTTTTTTTTTCAGTAACCTGGGTATATGTTTT	1467		
244	GTTTGTTCGGTCTCTGTAAAAAACCCTGGATTTTTTTTTTTCAGTAACCTGGGTATATGTTTT	303		
1468	CTCTTAAATAAGGTAATGAATGGCTTCGCCACAAATTTACCTTGACTACGATATCATCG	1527		
304	CTCTTAAATAAGGTAATGAATGGCTTCGCCACAAATTTACCTTGACTACGATATCATCG	363		
1528	ACATGACTTCTCTCAAAAAAAGAAATGCTTCATAGTTGTATTTTAAATTTGATATGTGAA	1587		
364	ACATGACTTCTCTCAAAAAAAGAAATGCTTCATAGTTGTATTTTAAATTTGATATGTGAA	423		
1588	AGAGTCATATTTTCCAGTTTATATTTCTAAGAAGAAGATAGATCATATAATCTGACAAG	1647		
424	AGAGTCATATTTTCCAGTTTATATTTCTAAGAAGAAGATAGATCATATAATCTGACAAG	483		
1648	GAAAAAGTTGCTTACCCCAAAATCTAAGTCTCAATCCCTGAGCCTCAGCAAAAACAGCTCC	1707		
484	GAAAAAGTTGCTTACCCCAAAATCTAAGTCTCAATCCCTGAGCCTCAGCAAAAACAGCTCC	1733		
1708	CTCCGAGGGAATCTTATACHTTAT	1733		
544	CCCCCTCCGAAGGGAAAAATCCCTTAAT	569		

RESULT 11
AA868507/c
LOCUS
DEFINITION
3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
EST.
AA868507.1 GI:2963952
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 504)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio1.llnl.gov/bbrp/image/image.html
Insert Length: 667 Std Error: 0.00

609 CATCATTTACTTCGATCAGATCCTGGTGAATGGTAAATTTTTCACATTTGGAGTCTGT 668
 Db |||||
 184 CATCATTTACTTTGATCAGATCCTGGTGAATGGTAAATTTTTCACATTTGGAGTCTGT 243
 QY |||||
 669 CTTTGTAGCACAGAAAGAAATTTACAGTTTCAGTTTTCAGTGATTAAGTCTACCA 728
 Db |||||
 244 CTTTGTGGCACCGAGGAAGAAATCTATAGTTTCAGTTTTCACGTAATTAAGTCTACCA 303
 QY |||||
 729 GAGCCAAATATCCAGGTTAACTTGATGTTAAATGGAACCCAGTAAATATCTGCTTGC 788
 Db |||||
 304 GAGCCAAACCATCCAGGTTAACTTGATGTTAAATGGAACCCAGTAAATATCTGCTTGC 363
 QY |||||
 789 GGGGGAACAAGATCTTACTGCTGAAGCTGCCAGATGCTGCTGCTTACCTAGATATA 848
 Db |||||
 364 TGGTGATAAGATGTGACCCGCTGAAGCCGCACTAATGGAGTGTCTCTGTACCTGGCAA 423
 QY |||||
 849 AGAGGATAAGTTTACCTTAACCTGGAAGAAAGTAAATTTGTTGGAGCTGGCGAGTATTC 908
 Db |||||
 424 AGAAGATAAGTCTACCTTAACCTGGAAGAAAGTAAATTTGTTGGAGCTGGCGAGTATTC 483
 QY |||||
 909 CACGTTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 968
 Db |||||
 484 CACGTTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 543
 QY |||||
 969 CAGTGAGGATGATCCACCTCTGAGTATTTGGAAGATCAATTTTTCATCATTTGGATTA 1028
 Db |||||
 544 AGTGAGGCGG-----CCACCCCGGTTATCAGAAGAGCACTTTGGAATGATGCTCTT 598
 QY |||||
 1029 TGTCTTTTATGTTTCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
 Db |||||
 599 ACTGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
 QY |||||
 1088 CCAATACAAATTTTCACAGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
 Db |||||
 656 CAGACAAAGTTTCACAGATCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701

RESULT 14

BU781883 436 bp mRNA linear EST 11-OCT-2002
 LOCUS im96902.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123050
 DEFINITION 5' similar to SW:CEHL_RAT P98087 CEREBELLIN-LIKE GLYCOPROTEIN. i,
 mRNA sequence.
 BU781883
 BU781883.1 GI:23824598
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 436)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V.,
 Williams, T., Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biochem.harvard.edu
 Library was constructed by Dr. J. Ferrer In vivo mass-excised to
 pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
 University Genome Sequencing Center For information on obtaining a
 clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 302.

FEATURES

source

Location/Qualifiers
 1..436
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IWAGB:6123050"
 /tissue_type="insulinoma"
 /lab_hosts="DH10B (phage-resistant)"
 /clone_lib="Human insulinoma"
 /notes="Organ: pancreas; Vector: pBluescript SK-; Site 1:
 XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
 pBluescript SK- by Dr. H. Inoue following the Washington
 University protocol
 (http://genome.wustl.edu/est/lambda_protocol.shtml).
 Please contact Hiroshi Inoue, MD/PhD for further
 information on this library (Metabolism Division, Permutt
 Laboratory, Washington University School of Medicine, Box
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
 is a Washington University Pancreas EST project library."

ORIGIN

Query Match 19.4%; Score 399.8; DB 13; Length 436;
 Best Local Similarity 94.9%; Pred. No. 1.8e-61;
 Matches 413; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
 QY 127 CAAGGTAAACAGAGCCAGAGAGCTCGAGGGAGACTTTGACTTCAAGCCACAGAA 186
 Db |||||
 2 CACGAGTAAACAGAGCCAGAGAGCTCGAGGGAGACTTTGACTTCAAGCCACAGAA 61
 QY |||||
 187 TGGTGAAGTGTGCGCGCGCGCGCGCGCTCCTGACAGCGCTGTCGACCTAGCCGCT 246
 Db |||||
 62 TGGTGAAGTGTGCGCGCGCGCGCGCTCCTGACAGCGCTGTCGACCTAGCCGCT 121
 QY |||||
 247 AGCATCTTCCGAGACACCGGGATCCCGGGGTAGAGCGACACGCGGGGAGCACACAGCGCC 306
 Db |||||
 122 AGCATCTTCCGAGACACCGGGATCCCGGGGTAGAGCGAGCGCGGGGAGCACACAGCGCC 181
 QY |||||
 307 AGCGGCTGCGGTGCGGCACACCGGCTCACCATGGGCTCCGGGCGCGGGGCGCTGTCCGCG 366
 Db |||||
 182 AGCGGCTGCGGTGCGGCACACCGGCTCACCATGGGCTCCGGGCGCGGGGCGCTGTCCGCG 241
 QY |||||
 367 GTCCCGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
 Db |||||
 242 GTCCCGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
 QY |||||
 427 ACGAGCCCATCTGCTGAGGGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
 Db |||||
 302 ACGAGCCCATCTGCTGAGGGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
 QY |||||
 487 GACTCCAGGGCTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 546
 Db |||||
 362 GACTCCAGGGGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 421
 QY |||||
 547 GTCCGCTTCTCGGCG 561
 Db |||||
 422 GGGGGGCTTCTCGGCG 436

RESULT 15

BU382466

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 750)

BU382466 750 bp mRNA linear EST 28-NOV-2002
 LOCUS 603859208F1 CSEQCHN75 Gallus gallus cDNA clone CHEST86801 5', mRNA
 DEFINITION sequence.
 ACCESSION BU382466
 VERSION BU382466.1 GI:25890467
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallus.
 1 (bases 1 to 750).

AUTHORS

Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392

COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1. 750
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST86801"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN75"
/note="Organ: trunks; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query Match 19.2%; Score 396; DB 13; Length 750;
Best Local Similarity 76.5%; Pred. No. 8.8e-61;
Matches 578; Conservative 0; Mismatches 165; Indels 13; Gaps 7;
QY 572 CCAACACGAGCCATCGAGATGAGCAACAGACGCGCATCATTTCTCGATCAGATCC 631
Db 1 CCAACACGAGCCCTCGAGATGAGCAACAGACGCGCATCATTTCTCGACCATCC 60
QY 632 TGGTGAATGTGGGTAATTTTTCACATTCGAGTCTCTTTGTAGCACCAAGAAAGGAA 691
Db 61 TAGTAAACGTGGGCAATTTTTCACGTGGGAATCTGTCTTTGTCCACCAAGAAAGGAA 120
QY 692 TTACAGTTTCAGTTTTCAGTGAATTAAGTCTACAGAGCCAAACTATCCAGGTTAACT 751
Db 121 TTACAGCTTCAGTTTTCAGTGAATTAAGTCTACAGAGCCAAACATCCAGGTTAACT 180
QY 752 TGAATGTTAAATGGAAACAGGTAATATCTGCTTTTCGGGGGACAAAGATGTTACTCGTG 811
Db 181 TGATGCTCAATGGAAAGCCAGTCACTCTGCTTTTGTGGGACAGGACGTCACCTCGTG 240
QY 812 AAGTCCACGAGATGCTGCTCTACCTAGATAAGAGGATAAGGTTTACCTAAAC 871
Db 241 AAGTCCACGATACGAGGATCTGCTCTATCTAGCAAGGAGGATAAGGTTTACCTGAAGC 300
QY 872 TGGAGAAAGTAATTTGGTGGAGGCTGGCAGTATTCACGTTTCTCGCTTTCTGGCTTCTGGTGT 931
Db 301 TGGAGAAAGGTAATCTGGTTGGTGGATGGCAGTATTCAGTTTCTCGCTTTCTGGCTTCTGGTGT 360
QY 932 TCCCCCTATAGGATTCATTTCTCCATGATGTTCCAGGTGAGGATGACCCACTCCT 991
Db 361 TTCCCCCTGTA-AAGTCAATTTTCTGTGATGTTCTGTCAGTGTGGAGTCGTCATTGCT 419
QY 992 GAGTTATTGGAAGATCATTTTTCATTCATTTGATGATGTCCT-TTATTGGTTTCTCATG 1050

Db 420 CTGTCACATGAAGATCATCTCTGTCTATCGTGGGATTTGATGTTTCTTATTTGTTTTCATG 479
QY 1051 GGTGGATATGGATTC---TAAGGATTCCTAGCCTGTCTGAAACCAATACAAATTTTCACAG 1106
Db 480 GGTGAATATAGATTTCTCTGTATGGATTTTGGCCCCATCTGAACCTACACAGAGTTTCACA 539
QY 1107 ATATTTGTGTGTCTGTCTGTTCAGTATATTTGGATTGGGACTCTA--AGCAGATAATACC 1164
Db 540 GAATTCCTGTGTGT-TTAAATACATATATTTGGATTGAGACTAAAGCAGACGATAATAA 598
QY 1165 TATGCTTTAAATGTAACAGTCAAAAGCTGTCTGCAAGACTTATCTGAATTTCTATTTCTGTG 1224
Db 599 CTGTGCTTAATGTTTACCGTCTTAAAGCTGCTGACCGGTTTATTTCAG-GTATCGTATCTG 657
QY 1225 GGATTACTGAATTAGTTTACAGATGTGAAATTTTATTGTTTGTATTTTAAAGACTGGCAA 1284
Db 658 GGACTTATGGATTCATGAGAGAGTGA---TTTCCTTTTAACTGTGAAACGACTGGCAA 714
QY 1285 CCAGTCTTAAGGATTAGAAACTCTAAAGTTCTGAC 1320
Db 715 CCAGGTCCTATATTAAGCCAGAAATTTGAGTTCAGAC 750

Search completed: May 15, 2004, 06:48:05

Job time : 5405 secs